Accounting for Complex Structure in Diffusion Weighted Imaging Data using Volume Fraction Representations

Thijs DHOLLANDER

Dissertation presented in partial fulfillment
of the requirements for the degree of
Doctor in Engineering Science

Examination committee:

em. prof. dr. ir. W. Sansen, chair
prof. dr. ir. F. Maes, supervisor
prof. dr. S. Sunaert, co-supervisor
prof. dr. ir. P. Suetens
prof. dr. ir. D. Vandermeulen
prof. dr. P. Dupont
prof. dr. U. Himmelreich

prof. dr. D. C. Alexander
(University College London)
dr. J. D. Tournier
(King's College London)

April 28, 2014

Uitgegeven in eigen beheer, Thijs Dhollander, Kasteelpark Arenberg 10 bus 2441, B-3001 Heverlee (Belgium)

ISBN 978-94-6018-827-5

D/2014/7515/51

☺ “Disco Brains!”

Els Godecharle

My girlfriend, upon seeing a directionally encoded color tractogram for the first time.

And many times thereafter.

Voorwoord

Ik heb er lang over nagedacht – wat er nog maar eens toe geleid heeft dat ik dit weer op een onwezenlijk uur nog moet schrijven; dat moet ik toch dringend eens beginnen afleren – Engels of Nederlands, “preface” of “voorwoord”? Gezien de titel hierboven, is het dus duidelijk “voorwoord” geworden. Want hoe veel ik verderop in dit boekje ook te vertellen heb in het Engels, ik kan nog altijd het beste babbelen (en blijven babbelen) in het Nederlands. Wie mij kent, weet dat als ik begin, er geen eind aan komt. Ook hier is dat weerom niet anders, want ik ben alweer vele lijnen tekst verder zonder nog maar het minste met effectieve inhoud verteld te hebben. Nederlands dus! Al was het maar omdat ik de komende jaren noodgedwongen nog heel wat Engels ga (moeten) praten.

Wat je hier in handen hebt (of eventueel op je scherm bekijkt), is een doctoraat. Dat is een verkort Nederlands woord voor bloed, zweet en tranen, een straaltje hoop, een doorbraak, euforie, glorie, roem, en dan weer naar bloed, zweet en tranen. Dat alles herhaalt zich een paar keer, gespreid over een dikke 4 jaar. En dan schrijft de doctorandus een boek. Die laatste activiteit draait minder om de voorgenoemde euforie, glorie of roem; maar des te meer om het bloed, zweet en tranen. Gelieve het boekje dus niet uit te wringen; dat gaat *ferme plekken* maken. Maar het feit dat je het vast hebt, betekent dat ik het geschreven heb, eindelijk!

Niks is zo zalig als een voorwoord te schrijven (alles is relatief): niet alleen kan ik nog eens vlijtig doortypen in het Nederlands, ik heb daarenboven de vrijheid om niet elk woord 42 keer te moeten omdraaien, wikken, en wegen. Want dat heb ik de laatste jaren vaak gedaan. Mijn promotor, Frederik Maes, kan er van meespreken. Niet dat ik hem daar vaak actief mee heb lastig gevallen, maar het indirecte effect was er wel: Frederik, ik heb je – ondanks je terechte continue stimulansen – lang laten wachten op een eerste *journal paper*. Ik ben van nature een “omdraaier”. Niets is finaal als het niet 42 keer is omgedraaid, beredeneerd, geherinterpreteerd, en... tijd om deze zin hier toch maar eens te stoppen. Maar je gaf me altijd de vrijheid om (binnen ietwat redelijke grenzen) te doen wat ik

wou. Ik heb daar dus ook gretig gebruik van gemaakt. Ik ben me er ook van bewust dat ik jou, en ongetwijfeld menig ingenieur in het begeleidingscomité, een degelijk eindje uit de *comfort zone* moet gehaald hebben door het aan te durven gehele hoofdstukken van dit doctoraat te vullen met passages die – op z'n zachts uitgedrukt – wel wat afwijken van het gangbare formele taalgebruik. Je eerste reactie bij het ontvangen van de tekst was ook mooi gewikt en gewogen: *"Ik heb al gezien dat je hier en daar een wat informelere schrijfstijl hanteert (zoals je ook verantwoordt bij de inleiding van elk hoofdstuk). Hopelijk wordt dit door de lezers gesmaakt."* Maar weerom heb je het toegelaten, en dat apprecieerde ik altijd heel hard: zolang ik een ietwat redelijke uitleg had, heb je me steeds het vertrouwen gegund dat toeliet om de maximale beleving te halen uit die 4.5 jaar van dat doctoraat. Bedankt hiervoor!

Waar een promotor is, is ook een co-promotor (nu ja, niet persé, maar toch wel in mijn comité). En wat voor één! Ik weet niet of het door dat voorgenoemde onwezenlijk late uur komt, maar Stefan, ik weet *serieus* niet met welke woorden (maar daar maar superlatieven van) ik je kan omschrijven – als co-promotor én als mens. Aan jou heb ik mijn eerste ISMRM en bijhorende ervaringen te danken. Alleen daarvoor al ben ik je eeuwig dankbaar (Stockholm staat eeuwig in het geheugen gegrift!). Als je kon, heb je me altijd extra kansen en opportuniteiten gegeven. Van jou leerde ik dat wiskundige en politieke problemen wel eens vaker samenvallen. De inzichten die je me gaf om ietwat handig te functioneren in deze boeiende onderzoekswereld zijn op m'n beide handen veruit niet te tellen. Met eender welke praktische vragen kon ik bij jou terecht; maar het deed ook deugd dat je vaak voor belangrijke vragen een beroep kwam doen op mijn mening. Zelfs als ik er een zootje van maak, sta je daar nog steeds voor mij. Ongelooflijk.

Ook de rest van het begeleidingscomité – Paul, Dirk, Uwe en Patrick – mocht er zijn. Met veel geduld hebben ze (vaak lang) gewacht op mijn presentaties, verslagen en deze tekst. Maar altijd hebben ze de moeite genomen om met veel interesse zich te verdiepen in mijn bevindingen. Ik hoop dat het aan het einde van de rit ietwat duidelijk is geworden waar ik nu juist allemaal mee bezig was. Alvast bedankt dat jullie er altijd op vertrouwd hebben dat mijn werk en plannen weldegelijk iets plausibel voorstelden! Daarenboven, specifiek aan Patrick gericht: ook hartelijk bedankt om me via het postgraduaat in *advanced medical imaging* te laten kennis maken met deze boeiende wereld. Jij was de brug die me van een master in informatica naar een doctoraat over medische beeldverwerking heeft geloodst; iets waar ik nooit spijt van zal hebben.

Next up is a short switch to English to thank the rest of my jury members: Daniel and Donald, a big thanks goes to both of you for accepting to be part of the jury. You gave me your critical opinions and a nice (hard) time at the preliminary defense – just what I needed. ;-) I could really appreciate your

open and supportive attitudes towards the simple doctoral student that I am (was). Being at the academic level that you both are, I can imagine your daily schedules are anything but a wealth of free time; yet you both took the time to carefully read a whole book of my writings, with an amazing eye for detail. The corrections this allowed me to make, have nicely improved the quality of the final product!

Een jury is ook niks zonder voorzitter. Willy Sansen, je kwam er pas bij sinds de preliminaire verdediging; maar de rust en wijsheid die je uitstraalde, en de ordelijke structuur die je bracht in het gebeuren, zijn zeer welgekomen. Ik ben blij dat ik hiervoor op je kon/kan vertrouwen!

Een doctoraat is meer dan een comité en een jury alleen. Het is ook werken, en dus komen daar collega's bij kijken. Neen, geen exhaustieve oplijsting hier (handige truuk om geen namen te vergeten), maar gewoon wat *losse* namen en bedankjes. Allereerst natuurlijk aan iedereen: bedankt om hier in het *Louvre* ook te zitten, en te werken. Alleen zou het maar saai zijn, weet je wel. Maar dat is het hier nooit. Van in de vroege dagen toen bepaalde (hyperactieve) collega's het leven opvrolijkten, tot de recentere jaren waar... vreemd genoeg net hetzelfde gebeurt met elke volgende generatie collega's: het is hier altijd zo plezant dat zelfs de drukste deadlineperiode's niet als werken aanvoelen (maar vertel dat toch maar niet door aan de buitenwereld). Dan, namen dus (maar jullie zijn allemaal even veel waard uiteraard ;-)). Om te beginnen zijn er die 2 personen die verdacht veel opdoemen in het hieropvolgende academische CV. Tom en Daan, jullie zijn niet alleen heel goeie masterthesisstudenten gebleken, ook met het daaropvolgende eigen doctoraat zijn jullie stevig en goed bezig. Ik heb nooit veel "werk" aan jullie gehad, en ben blij dat jullie snel zelf uitgegroeid zijn tot heel competente en kritische onderzoekers. Janaki, you were my neighbor for quite some years. Thanks for the many good conversations we had, and all your advice to get me started in the beginning, for my first conference, and even in the end when I was writing this book. Jeroen, jouw kritische geest heeft de groep altijd scherp gehouden; maar vooral veel karakter gegeven. Bart en Dominique, bedankt voor alle technische ondersteuning! Annitta en Patricia, bedankt voor alle administratieve ondersteuning en inzichten in de bijhorende jungle! Jose and David, thanks for being my neighbors during the last few months! Catarina, thanks for bringing some authentic Portuguese warmth to our group (and *teaching* me a thing or two about ultrasound along the way). Louise, Thibo, Sofie, Sabine, ... radiologie, bedankt om de nuttige andere helft van het multi-disciplinaire domein te vertegenwoordigen en ons van wat praktisch inzicht te voorzien. Iedereen: nogmaals bedankt, om er te zijn en te maken van de groep wat die is (schitterend!). Ik ga dat hier nog *fameus* missen denk ik. Ook dank aan Wim, Ben, Jelle, de overige collega's uit Antwerpen, én uit Nederland, *and all the other colleagues worldwide*, om me op weg te zetten;

en van de vele conferenties telkens weer memorabele gebeurtenissen te maken!

De laatste jaren waren meer dan werken alleen natuurlijk. Velen weten het, of niet, maar ik vul een stukje van m'n vrije tijd met het beoefenen van tai chi (diverse *kung fu geluiden* hier vrij in te voegen). Dat ik hiermee ooit heb kennis gemaakt, heb ik vooreerst aan Peter te danken: Peter, bedankt om ooit met veel moed voor telkens nieuwe mensen de basics van tai chi proberen uit te leggen en er iets leuk van te maken. Het echte *trainen* begon dan weer bij m'n leraar: Fred, bedankt om ons zo hard te laten werken. Door het uitgebreid herhalen meermaals per week, kan ik nu niet alleen trots tegen iedereen zeggen "dat ik tai chi kan" (en zo vele vreemde blikken opwekken), maar tenminste ook onderbouwen dat het een gevechtskunst is (dat komt toch altijd weer wat *stoerder* over) en dat we – als we daar zin in hebben – mekaar in het rond gooien, klemmen, bijna breken, wurgen, zwaard-, stok- en speergevechten aangaan; maar dat alles met de nodige zin voor verantwoordelijkheid en een occasionele blauwe plek. Last but not least: thank you Master Shen Tiegen! Not only have you showed eternal patience in teaching us even the most basic techniques; you also proved that many things are simply natural. Your simple but powerful advice "Change!" provides an answer to many things, even far beyond the practice of tai chi: it works amazingly well in many daily activities, and even in the competitive academic world. Naast de leraars, moeten evenzeer de mede-studenten bedankt worden. Het is een plezier om samen te trainen, en van elkaar steeds weer nieuwe dingen te leren!

Ook nog maar eens dank aan de "mannen van *Sinnekloas*", al een constante sinds de middelbare school, doorheen de studententijd en nu terwijl we allemaal (semi-) werkmensen geworden zijn. Mannen, jullie zijn altijd schitterende vrienden gebleven; en de beste reden om met regelmaat *veirkesrebbekes* in Het Laatste Avondmaal te gaan eten en nog eens bij te praten.

Met lichte overlap met de vorige groep, al evenzeer dank aan de vriendengroep uit de informatica jaren, voortgevloeid uit het legendarische pémé-groepje "de *partypoopers*" (ik ga echt moeite hebben om het taalgebruik hierna nog terug recht te trekken; laat staan ooit nog professioneel over te komen bij al m'n collega's die dit lezen). Ook jullie zijn een constante gebleven die ongetwijfeld voor het leven meegaat. En hoe *nerderig* ik ook nog steeds kan zijn hier bij de collega's, er gaat uiteraard niets boven een onderonsje met rasechte informatici. Met alle risico's van dien (professionele geloofwaardigheid gaat op dit exacte punt drastisch onder nul!), ook hartelijk bedankt om me ooit die paar keer weer veilig naar kot te brengen/ondersteunen (laat ons hopen dat het niet dragen was...?). ...of hoe we ooit student waren nog ruim voor de jeugd daar de term "YOLO" voor nodig had. We weten allemaal dat het beter was *in onzen tijd*.

Hoewel ik nu officieel dreig een diploma te halen waar het woord "ingenieur" op

voorkomt, ben ik ook nog steeds trots op m'n jaren bij Wina! Bedankt aan elke Winees om mede te zorgen voor een schitterende studententijd. Ook dank aan elk mede-presidiumlid dat ooit de revue is gepasseerd: al die jaren presidium waren een zeer leerrijke ervaring. Dat ik als preses van Wina ooit openlijk in het studentenblad Veto heb verkondigd dat toegepaste wetenschap iets heel vies is, gaat me nu met veel plezier nog levenslang achtervolgen vrees ik. :-)

Ook wonen hebben we de laatste jaren gedaan; en wel als huurders bij Frans, Maria en Jo. Deze wereld draait rond omdat er nog degelijke vakmensen en huisvaders en -moeders op rondlopen, en daar zijn de voorgenoemden het beste voorbeeld van. Niet alleen hebben ze ons voorzien van een mooi plaatsje om te wonen, maar ook nog eens van de occasionele koffie, wafels, verse eitjes, gouden raad en leerrijke gesprekken. Ze zorgen voor hun huurders als waren het hun eigen kinderen. Altijd vriendelijk en opgewekt; een toonbeeld van levensvreugde en geluk!

Back to English! Did I mention that I will be moving to Australia? Apart from everyone who brought me up to this point in my career, a very special thanks goes out to David Raffelt. Not only was more than half of this PhD inspired by his established works; he's also the guy that informed me of opportunities at the Florey Institute of Neuroscience and Mental Health, down under! Thanks Dave, you are officially the original cause of my next big adventure (not to mention, a future colleague!). Another important part of this, is the fact that there actually exists someone on the other side of this planet that is willing to pay me to perform science over there; and that would be Alan Connelly. I'm already amazed by the efforts you are taking to even get me over there in the first place, Alan. I can't wait to actually get over there and start doing some seriously interesting work. Thanks again for putting so much trust in me!

Terwijl ik al de voorgaande mensen – en verder iedereen die me in dit leven tot op dit punt heeft gebracht of gelijk wanneer de revue is gepasseerd – nogmaals uitgebreid wil bedanken voor hun rol, zijn er een paar mensen die een extra vermelding verdienen; of maak daar maar gerust een gouden standbeeld van.

Mama, bedankt voor alles! Zonder de kansen die je me altijd hebt gegeven, had ik dit nooit kunnen bereiken. Binnenkort ga ik echter echt wel zelf m'n was moeten beginnen doen, maar ik beloof je nu alvast dat ik uitermate voorzichtig zal zijn (en je mag het me eerst zélf aanleren voor de veiligheid :-)). En ik heb natuurlijk Els bij me om een oogje in het zeil te houden!

Marleen en John, bedankt om me zowel te vertrouwen met jullie dochter alsook om me op te nemen in de familie! Bij uitbreiding ook bedankt aan heel die familie om dat al even hartelijk te doen. En aan Bonneke beloof ik bij deze dat ik heel goed zorg zal dragen voor "ons Els"! ;-)

Ziezo, nog mooi een volledige pagina over. Drie-twee-één verrassing: ik heb een vriendin! Niet alleen siert ze een van de eerdere pagina's in dit boekje met een professioneel filosofische quote, ze krijgt er ook nog eens deze pagina bij. Op de dag dat dit boekje *en masse* wordt uitgedeeld, zullen we het al zo'n 6 jaar, 4 maanden en 24 dagen bij elkaar hebben uitgehouden. Dat is omdat 6×4 gelijk is aan 24, en omdat alle planeten in zowel dit als de 24 meest nabije sterrenstelsels op 1 lijn stonden op 4 december in het jaar 2007. Ze is zo mooi als de hemel op... een hele mooie dag, danst als een prinsesje en zingt vrolijker "diep in de zee" dan Sebastiaan de krab. Dat laatste onder meer als spontane reactie op Fig. 4.10 en Fig. 5.2 in dit doctoraat. Want die hersenen lijken toch zo sprekend op een zwerm kwallen. Achter elke man staat een sterke vrouw, wordt wel eens gezegd. Ik hoop dat die mannen daar mee kunnen leven, want ik heb iets veel beters: een sterk Elsje! Er valt al iets te zeggen voor zo'n knap meisje dat zich spontaan wil binden aan een informaticus, en daar ook nog mee kan leven als die een doctoraat in de ingenieurswetenschappen maakt; maar dat ze het ook nog eens ziet zitten om ermee "zowat in Australië" te gaan leven, maakt er wel zeker eentje van om te houden. Ik zou er nu al spontaan voor op m'n knie gaan, ware het niet dat het laatste mopje dat ik in die context ooit heb gemaakt op een *klein paniekske* uitdraaide. Gelukkig hebben we over de jaren heen ons gevoel voor humor weten gelijk te stemmen (lees: ze heeft nu bijna een master in internetcultuur en de bijhorende humor). Samen hebben we ook al dolle avonturen mogen beleven; met als hoogtepunt het vast komen te zitten op een klein eilandje voor de Siciliaanse kust, daar verbranden van de zon en ons *onnozel verschieten* van een slang. En dus zullen we ook probleemloos overleven in Australië.

Maar laten we ook nog even terug reflecteren op het eerder genoemde bloed, zweet en tranen. Zie, daar bewijst zo'n kwaliteitsvolle aanschaf zijn waarde dan weer meer dan eens te voren. Sta me toe hier heel duidelijk zijn: de laatste maanden waren afzien. Het valt niet aan te raden om nog snel-snel een paper te schrijven op een paar maand, tesamen met een hele boek, tesamen met het proberen verzamelen van alle materiaal voor een visumaanvraag, tesamen met leren autorijden (waarom ook niet?), tesamen met te proberen eten en slapen. Elsje, m'n allerliefste *schattigaard* (minpunten voor deze term, maar daar kunnen we mee leven ;-)), bedankt om me door dik en dun te steunen! Niks in deze hele wereld klinkt beter dan jij die vrolijk "diep in de zee[1]" begint te zingen als ik absoluut geradbraakt thuis kom na een dag of nacht werken. En daarom heb je nu de volle 2 pagina's gewonnen in m'n doctoraat!

Thijs Dhollander

[1] ...al die sardientjes, zijn ook mijn vriendjes, JIPPIE-JA-JEEEEE!

Academic CV

Thijs Dhollander

thijs.dhollander@gmail.com

Education

- **Ph.D., Engineering Science, 2014**
  KU Leuven, Leuven, Belgium
  Dissertation: "Accounting for complex structure in diffusion weighted imaging data using volume fraction representations"
- **Postgraduate Studies in Advanced Medical Imaging, 2009**
  KU Leuven, Leuven, Belgium
  Dissertation: "A framework for graph-theoretical analysis of functional brain networks"
- **M.Sc., Informatics, 2008**
  KU Leuven, Leuven, Belgium
  Dissertation: "Matchingproblemen voor vlakke grafen" ("Matching problems for planar graphs")
- **General High School, Sciences-Mathematics, 2003**
  Sint-Jozef-Klein-Seminarie, Sint-Niklaas, Belgium

Awards

- **Summa Cum Laude Merit Award (top 3% of abstracts)** for "Robustifying Probabilistic Tractography by using Track Orientation Distributions." (21st annual meeting of the ISMRM; Salt Lake City (United States), 2013)
- **Certificate of Merit Award (top 10 of abstracts)** for "How reliable are findings from Track Density Imaging?" (29th annual meeting of the ESMRMB; Lisbon (Portugal), 2012)
- **Summa Cum Laude Merit Award (top 3% of abstracts)** for "Track-density Imaging & Noise: when Super-resolution Quality does not yield Accuracy." (20th annual meeting of the ISMRM; Melbourne (Australia), 2012)
- **Nominated for the Diffusion Study Group Traditional Poster Award (top 4)** for "Track-density Imaging & Noise: when Super-resolution Quality does not yield Accuracy." (20th annual meeting of the ISMRM; Melbourne (Australia), 2012)
- **Nominated for the Best Poster Award (top 8)** for "Track-density Imaging + Noise: when Super-resolution Quality does not yield Accuracy." (4th annual meeting of the ISMRM Benelux Chapter; Leuven (Belgium), 2012)
- **Best Presentation Award (1st place)** for "Constructing a Hybrid Diffusion Imaging Atlas in Q-space." (3th annual meeting of the ISMRM Benelux Chapter; Hoeven (Netherlands), 2011)

Presentations

Oral (conference)

- Robustifying Probabilistic Tractography by using Track Orientation Distributions. (21st annual meeting of the ISMRM; Salt Lake City (United States), 2013)
- Using Track Orientation Distributions to Robustify Probabilistic Tractography. (5th annual meeting of the ISMRM Benelux Chapter; Rotterdam (Netherlands), 2013)

- Exploring how I connect with my Mind: Diffusion Weighted Imaging of the brain. (iMinds the conference; Ghent (Belgium), 2012)
- How reliable are findings from Track Density Imaging? (29th annual meeting of the ESMRMB; Lisbon (Portugal), 2012)
- A Hybrid Diffusion Imaging Atlas in Q-space. (replacing/invited talk for CDMRI workshop at 14th international conference on MICCAI; Toronto (Canada), 2011)
- A Hybrid Diffusion Imaging Atlas in Q-space. (19th annual meeting of the ISMRM; Montreal (Canada), 2011)
- Constructing a Hybrid Diffusion Imaging Atlas in Q-space. (3th annual meeting of the ISMRM Benelux Chapter; Hoeven (Netherlands), 2011)
- Spatial Transformations of High Angular Resolution Diffusion Imaging Data in Q-space. (CDMRI workshop at 13th international conference on MICCAI; Beijing (China), 2010)

Poster (conference)

- Quantitative Validation of TOD-based Tractography by a Tractometer Approach. (ISMRM Workshop on Diffusion as a Probe of Neural Tissue Microstructure; Podstrana (Croatia), 2013)
- DTI vs. higher-order models. (diffusion study group double poster debate at 21st annual meeting of the ISMRM; Salt Lake City (United States), 2013)
- Track-density Imaging & Noise: when Super-resolution Quality does not yield Accuracy. (20th annual meeting of the ISMRM; Melbourne (Australia), 2012)
- Track-density Imaging + Noise: when Super-resolution Quality does not yield Accuracy. (4th annual meeting of the ISMRM Benelux Chapter; Leuven (Belgium), 2012)
- Feasibility and Advantages of Diffusion Weighted Imaging Atlas Construction in Q-space. (14th international conference on MICCAI; Toronto (Canada), 2011)
- Methods for Reorienting and Retransforming Diffusion Weighted Imaging Data. (19th annual meeting of the ISMRM; Montreal (Canada), 2011)

Research seminars

- Diffusion Weighted Imaging (DWI), Tractography, Track Density Imaging (TDI) and Track Orientation Distributions (TODs): a high-level overview. (KU Leuven, 2013)
- Track-density imaging (TDI): super-resolution magic, illusion, or just a clever trick? (KU Leuven, 2012)
- Spatial Transformations of High Angular Resolution Diffusion Imaging Data in Q-space. (KU Leuven, 2010)
- Diffusion Weighted Imaging: Spatial Transformations. (KU Leuven, 2010)
- Diffusion Weighted Imaging: Terminology & Models. (KU Leuven, 2009)

Publications

Journal

- **Dhollander, T.**, Jeurissen, B., Maes, F., Sunaert, S., Suetens, P., Tournier, J.D.: Multi-Shell Multi-Tissue (MSMT) representation and Preservation of Principal Volume Fractions (PPVF) retransformation of DWI signal. (in preparation)
- Jeurissen, B., **Dhollander, T.**, Tournier, J.D., Connelly, A., Sijbers, J.: Constrained spherical deconvolution of multi-shell diffusion MRI data. (in preparation)
- Nelissen, N., Wang, Y., **Dhollander, T.**, Adamczuk, K., Dupont, P., Vandenberghe, R.: Graph analysis of structural connectivity and relation to functional connectivity within the associative-semantic network. (in preparation)
- **Dhollander, T.**, Emsell, L., Van Hecke, W., Maes, F., Sunaert, S., Suetens, P.: Track Orientation Density Imaging (TODI) and Track Orientation Distribution (TOD) based tractography. NeuroImage (in press) (2014)
- Heitger, M.H., Goble, D.J., **Dhollander, T.**, Dupont, P., Caeyenberghs, K., Leemans, A., Sunaert, S., Swinnen, S.P.: Bimanual Motor

Coordination in Older Adults is Associated with Increased Functional Brain Connectivity – A Graph-Theoretical Analysis. PLoS ONE 8(4), e62133 (2013)

- Vandenberghe, R., Wang, Y., Nelissen, N., Vandenbulcke, M., **Dhollander, T.**, Sunaert, S., Dupont, P.: The associative-semantic network for words and pictures: Effective connectivity and graph analysis. Brain and Language (in press) (2012)
- Heitger, M.H., Ronsse, R., **Dhollander, T.**, Dupont, P., Caeyenberghs, K., Swinnen, S.P.: Motor learning-induced changes in functional brain connectivity as revealed by means of graph-theoretical network analysis. NeuroImage 61(3), 633–650 (2012)
- Caeyenberghs. K., Leemans, A., Heitger, M.H., Leunissen, I., **Dhollander, T.**, Sunaert, S., Dupont, P., Swinnen, S.P.: Graph analysis of functional brain networks for cognitive control of action in traumatic brain injury. Brain 135(4), 1293–1307 (2012)

Bookchapter

- **Dhollander, T.**: From diffusion to the diffusion tensor. In: Van Hecke, W., Emsell, L., Sunaert, S. (eds.) Practical Handbook of DTI. To be published by Springer, Heidelberg.

Conference (proceedings/abstract)

- Jeurissen, B., Tournier, J.D., **Dhollander, T.**, Connelly, A., Sijbers, J.: Brain tissue types resolved using spherical deconvolution of multi-shell diffusion MRI data. ISMRM 22 (in press) (2014)
- Emsell, L., **Dhollander, T.**, Nelissen, K., Peeters, R., Vanduffel, W., Sunaert, S.: Visualizing complex white matter anatomy in the live monkey at 3T using super-resolution track density imaging. ISMRM 22 (in press) (2014)
- Billiet, T., Deprez, S., Maedler, B., Peeters, R., Zhang, H., Leemans, A., **Dhollander, T.**, Christiaens, D., Amant, F., Smeets, A., Van den Bergh, B., Vandenbulcke, M., Legius, E., Sunaert, S., Emsell, L.: Investigating the long-term effects of systemic chemotherapy on brain white matter using multi-shell diffusion MRI and myelin water imaging. ISMRM 22 (in press) (2014)

- Billiet, T., Deprez, S., Maedler, B., Peeters, R., Zhang, H., Leemans, A., **Dhollander, T.**, Christiaens, D., Amant, F., Smeets, A., Van den Bergh, B., Vandenbulcke, M., Legius, E., Sunaert, S., Emsell, L.: Investigating the long-term effects of systemic chemotherapy on brain white matter using multi-shell diffusion MRI and myelin water imaging. ISMRM Benelux Chapter 6, 54 (2014)
- **Dhollander, T.**, Emsell, L., Van Hecke, W., Maes, F., Sunaert, S., Suetens, P.: Quantitative Validation of TOD-based Tractography by a Tractometer Approach. ISMRM Workshop on Diffusion as a Probe of Neural Tissue Microstructure, 90 (2013)
- Christiaens, D., **Dhollander, T.**, Maes, F., Sunaert, S., Suetens, P.: Groupwise Deformable Registration of Fiber Track Sets using Track Orientation Distributions. MICCAI 16, CDMRI Workshop, pp. 149–159 (2013)
- **Dhollander, T.**, Emsell, L., Van Hecke, W., Maes, F., Sunaert, S., Suetens, P.: Robustifying Probabilistic Tractography by using Track Orientation Distributions. ISMRM 21, 774 (2013)
- Haeck, T., **Dhollander, T.**, Maes, F., Sunaert, S., Suetens, P.: Feasibility of atlas-based segmentation of the brain in the presence of tumor by a weighted least-squares demons algorithm. ISMRM 21, 1242 (2013)
- **Dhollander, T.**, Emsell, L., Van Hecke, W., Maes, F., Sunaert, S., Suetens, P.: Using Track Orientation Distributions to Robustify Probabilistic Tractography. ISMRM Benelux Chapter 5, 56 (2013)
- Haeck, T., **Dhollander, T.**, Maes, F., Sunaert, S., Suetens, P.: Feasibility of Tumor Robust Atlas-based Segmentation of the Brain by a Weighted Least-squares Demons Algorithm. ISMRM Benelux Chapter 5, 92 (2013)
- Christiaens, D., **Dhollander, T.**, Maes, F., Sunaert, S., Suetens, P.: On the Estimation of the Fiber Response Function for Constrained Spherical Deconvolution. ISMRM Benelux Chapter 5, 106 (2013)
- **Dhollander, T.**, Emsell, L., Van Hecke, W., Maes, F., Sunaert, S., Suetens, P.: How reliable are findings from Track Density Imaging? ESMRMB 29, 240 (2012)
- Billiet, T., **Dhollander, T.**, Peeters, R., Deprez, S., Mädler, B., Sunaert, S., Emsell, L.: Quality assurance of multi-echo data used for myelin water mapping through residual bootstrapping. ESMRMB 29, 243 (2012)

- Christiaens, D., **Dhollander, T.**, Maes, F., Sunaert, S., Suetens, P.: The Effect of Reorientation of the Fibre Orientation Distribution on Fibre Tracking. MICCAI 15, CDMRI Workshop, pp. 33–44 (2012)
- **Dhollander, T.**, Emsell, L., Van Hecke, W., Maes, F., Sunaert, S., Suetens, P.: Track-density Imaging & Noise: when Super-resolution Quality does not yield Accuracy. ISMRM 20, 1920 (2012)
- **Dhollander, T.**, Emsell, L., Van Hecke, W., Maes, F., Sunaert, S., Suetens, P.: Track-density Imaging + Noise: when Super-resolution Quality does not yield Accuracy. ISMRM Benelux Chapter 4, 65 (2012)
- **Dhollander, T.**, Veraart, J., Van Hecke, W., Maes, F., Sunaert, S., Sijbers, J., Suetens, P.: Feasibility and Advantages of Diffusion Weighted Imaging Atlas Construction in Q-space. In: Fichtinger, G., Martel, A., Peters, T. (eds.) MICCAI 2011, Part II, LNCS vol. 6892, pp. 166–173. Springer, Heidelberg (2011)
- Haeck, T., **Dhollander, T.**, Van Hecke, W., Maes, F., Sunaert, S., Suetens, P.: Interpolation of High Angular Resolution Diffusion Imaging Data by Imposing Distances on Q-space. MICCAI 14, CDMRI Workshop, pp. 64–75 (2011)
- Heitger, M.H., Goble, D.J., **Dhollander, T.**, Dupont, P., Swinnen, S.P.: Graph-theoretical analysis of functional brain connectivity during bimanual coordination in the elderly. Proceedings of the 17th Annual Meeting of the Organization for Human Brain Mapping, 2815 (2011)
- Caeyenberghs, K., Heitger, M.H., Leunissen, I., Geurts, M., **Dhollander, T.**, Dupont, P., Swinnen, S.P.: Graph analysis of functional brain networks for cognitive control in traumatic brain injury. Proceedings of the 17th Annual Meeting of the Organization for Human Brain Mapping, 2489 (2011)
- **Dhollander, T.**, Van Hecke, W., Maes, F., Sunaert, S., Suetens, P.: A Hybrid Diffusion Imaging Atlas in Q-space. ISMRM 19, 416 (2011)
- **Dhollander, T.**, Van Hecke, W., Maes, F., Sunaert, S., Suetens, P.: Methods for Reorienting and Retransforming Diffusion Weighted Imaging Data. ISMRM 19, 4664 (2011)
- **Dhollander, T.**, Van Hecke, W., Maes, F., Sunaert, S., Suetens, P.: Constructing a Hybrid Diffusion Imaging Atlas in Q-space. ISMRM Benelux Chapter 3, 43 (2011)

- **Dhollander, T.**, Van Hecke, W., Maes, F., Sunaert, S., Suetens, P.: Spatial Transformations of High Angular Resolution Diffusion Imaging Data in Q-space. MICCAI 13, CDMRI Workshop, pp. 73–83 (2010)
- Heitger, M.H., Ronsse, R., **Dhollander, T.**, Dupont, P., Wenderoth, N., Swinnen, S.P.: Graph-theoretical network analysis shows changes of functional brain connectivity in motor learning. Proceedings of the 16th Annual Meeting of the Organization for Human Brain Mapping, 728 (2010)

Experience

- Ombuds person for the "Postgraduate Studies in Advanced Medical Imaging" program (KU Leuven, 2010–2013)
- Teaching assistant (exercise sessions) "Digital Signal Processing" (KU Leuven, 2009–2013)
- Daily supervisor/mentor of master thesis students:
  - Tom Haeck (KU Leuven, 2010–2011)
  - Daan Christiaens (KU Leuven, 2011–2012)
  - Lucas Koelman (KU Leuven, 2012–2013)
  - Xavier Garcia Pich (KU Leuven, 2013–2014) (co-supervised by Tom Haeck)
  - Dorothée Vercruysse (KU Leuven, 2013–2014) (co-supervised by Daan Christiaens)

Language skills

- Dutch: native
- English: fluent
- French: fair

Abstract

The domain of diffusion weighted imaging (DWI) has come a long way since its initial development in the mid-1980s. Over the years, we have gained a better understanding of the accompanying techniques and necessary processing steps involved, furthermore leading to a wealth of new insights in the complex workings of the (human) brain. The introduction of diffusion tensor imaging (DTI) has played a crucial role in this process, as it provided the first model intended to deal with anisotropic diffusion; a particular feature observed in the white matter (WM), as opposed to the other most common "tissue" types found in the human brain, *i.e.* the gray matter (GM) and cerebrospinal fluid (CSF).

During the last decade, however, we have come to realize that the DTI model is severely lacking in its possibilities to represent voxels that contain so called "crossing fibers", a general name that is often used to refer to a range of complex geometric fiber configurations caused by the partial volume effect.

In this PhD thesis, we intended to design certain representations of (information extracted from) DWI data that take into account the aforementioned variety of complex geometrical configurations. Our proposed novel representations aim to offer a greater flexibility that should inherently render many existing difficult problems (*e.g.* segmentation and registration) trivial; yet make as little assumptions as possible on the nature of the data or the properties of the underlying structures.

A first major contribution is a generic framework for multi-shell multi-tissue (MSMT) representations, and a specific implementation tailored to represent WM, GM and CSF in the human brain. This representation was specifically designed to render the retransformation problem trivial. The latter was easily solved by a newly introduced preservation of principal volume fractions (PPVF) retransformation strategy.

A second major contribution is a more tangible track orientation distribution (TOD) representation for complex fiber track distributions. Our newly developed

method to obtain such a TOD, is termed track orientation density imaging (TODI). This technique allowed us to gain further understanding in the amplitude of a short-tracks TOD, which can now be interpreted as a measure of track-like local support (TLS). Furthermore, we showed that employing the latter for TOD-based tractography results in guiding the tracks along directions that are more likely to correspond to continuous structure over a longer distance; *i.e.* track-like structure!

Both representations (MSMT and the TOD) forthcoming from these major contributions, also allow for increased insight in many other aspects of the data they describe, and provide a large range of opportunities for future research.

Samenvatting

Het domein van diffusie gewogen beeldvorming (diffusion weighted imaging, DWI) is ver geëvolueerd sinds zijn initiële ontwikkeling in het midden van de jaren 1980. Doorheen de tijd hebben we betere inzichten verworven in de bijhorende technieken en noodzakelijke verwerkingsstappen die hierbij betrokken zijn, wat op zich geleid heeft tot een rijkdom aan nieuwe inzichten in de ingewikkelde werking van de (menselijke) hersenen. De introductie van diffusie tensor beeldvorming (diffusion tensor imaging, DTI) heeft een cruciale rol gespeeld in dit proces, aangezien dit het eerste model leverde dat bewust omging met anisotrope diffusie; een specifieke eigenschap waargenomen in de witte stof (white matter, WM), in tegenstelling tot de andere gangbare "weefsel" types in de menselijke hersenen, *i.e.* de grijze stof (gray matter, GM) en de cerebrospinale vloeistof (cerebrospinal fluid, CSF).

Gedurende het laatste decennium, echter, zijn we er ons van bewust geworden dat het DTI model ernstig tekort schiet in zijn mogelijkheden om voxels voor te stellen die zogenaamde "kruisende vezels" bevatten; een algemene benaming typisch gebruikt om te verwijzen naar een waaier van complexe geometrische vezel configuraties veroorzaakt door het partiële volume effect.

In deze doctoraatsthesis was het onze bedoeling om bepaalde voorstellingen voor (informatie uit) DWI data te ontwikkelen, welke de voorgenoemde waaier aan complexe geometrische configuraties in rekening brengen. Onze nieuwe voorstellingen bieden een betere flexibiliteit aan die inherent verschillende bestaande problemen (*e.g.* segmentatie en registratie) triviaal maken; maar ze maken evenwel zo weinig mogelijk veronderstellingen over de aard van de data of de specifieke eigenschappen van de onderliggende structuren.

Een eerste grote bijdrage bestaat in een generisch framework voor multi-schil multi-weefsel (multi-shell multi-tissue, MSMT) voorstellingen, en een specifieke implementatie afgestemd op het voorstellen van WM, GM en CSF in de menselijke hersenen. Deze voorstelling is specifiek ontworpen om het

retransformatie probleem triviaal te maken. Dit probleem kon eenvoudig opgelost worden door een nieuwe behoud van hoofdzakelijke volume fracties (preservation of principal volume fractions, PPVF) retransformatie strategie.

Een tweede grote bijdrage bestaat in een meer tastbare track richtingsverdeling (track orientation distribution, TOD) voorstelling voor complexe vezel track verdelingen. Onze nieuw ontwikkelde methode om zo'n TOD te verkrijgen, heet track richtingdichtheids beeldvorming (track orientation density imaging, TODI). Deze techniek liet ons toe om verdere inzichten te verwerven in de amplitude van een korte-tracks TOD, welke nu kan geïnterpreteerd worden als een maat voor track-achtige lokale ondersteuning (track-like local support, TLS). Bovendien toonden we aan dat het gebruik van deze laatste voor TOD-gebaseerde tractografie leidt tot het sturen van tracks langsheen richtingen die met grotere waarschijnlijkheid overeenstemmen met continue structuur over een langere afstand; *i.e.* track-achtige structuur!

Beide voorstellingen (MSMT en de TOD) die voortkomen uit deze grote bijdrages, laten ook een verruimd inzicht toe in vele andere aspecten van de data die ze omschrijven, en leveren een grote waaier aan opportuniteiten voor toekomstig onderzoek.

Contents

Chapter 1

Brains and diffusion

1.1 The (human) brain

The book you're reading is all about brains, human ones mostly. For those not fully familiar with this marvelous yet intriguing organ, this section should provide the most important basic facts. Your brain is the central computing unit of your nervous system. It stores all your memories and knowledge, gets to process most of your input, and is in control of most of your output. Hence, it plays a massively important role, but yet, we still know relatively little about it.

Your brain is safely tucked away in your head. It is as such protected by several layers from the dangers of the outside world. On the outside, there's your skin (and possibly a nice set of hair). The skin is connected to your skull by an extra layer (the periosteum). The bone of the skull itself of course plays a crucial role towards the overall protection; it's your natural helmet. But it doesn't end there: further inwards we meet a collection of membranes, the meninges, that provide further protection: the dura mater (though and durable), the arachnoid mater (extra cushioning; it looks like a spider web) and finally the pia mater (soft, delicate and waterproof) firmly envelopping the outer surface of the brain iself. If you're normal and healthy, the skull, dura mater and arachnoid mater are firmly attached to each other, while the pia mater is attached to your brain. Between the arachnoid mater and the pia mater, there's a gap filled with cerebrospinal fluid. This provides important extra protection against some level of impact (otherwise, your soft brain might still slam against your rigid skull if you move your head). Apart from all this physical protection, there's also some important biological protection in place: the blood–brain barrier nicely controls

what can get through, and what cannot; a mechanism that should protect your brain from the serious threat of infection.

Behind all these barriers, we finally find your brain. If you're an adult, your brain should weigh in at a good 1.5 kg and comes as a nice volume of an average 1200 cm^3 smart stuff. Men's brains are slightly larger and heavier than women's; but this should not put the women at any cognitive disadvantages. Your brain itself consists mostly of two distinguished tissue types: the white matter and the gray matter. Fig. 1.1 provides an overal view on a coronal slice through the brain volume. Most of the gray matter is found on the outside of the brain; a layer, called the cortex, that features outward and inward foldings (respectively gyri and sulci) in order to maximize its surface area (measuring about 0.12 m^2). The cortex features an average thickness of about 2.5 mm. Underneath the cortex, we find a large mass of white matter. It's white because it contains a lot of myelin, which itself features a large amount of fat; but we'll get to that later. Finally, on the inside, there's a collection of cavities filled with cerebrospinal fluid, called the ventricles. This cerebrospinal fluid, by the way, consists mostly (99%) of water.

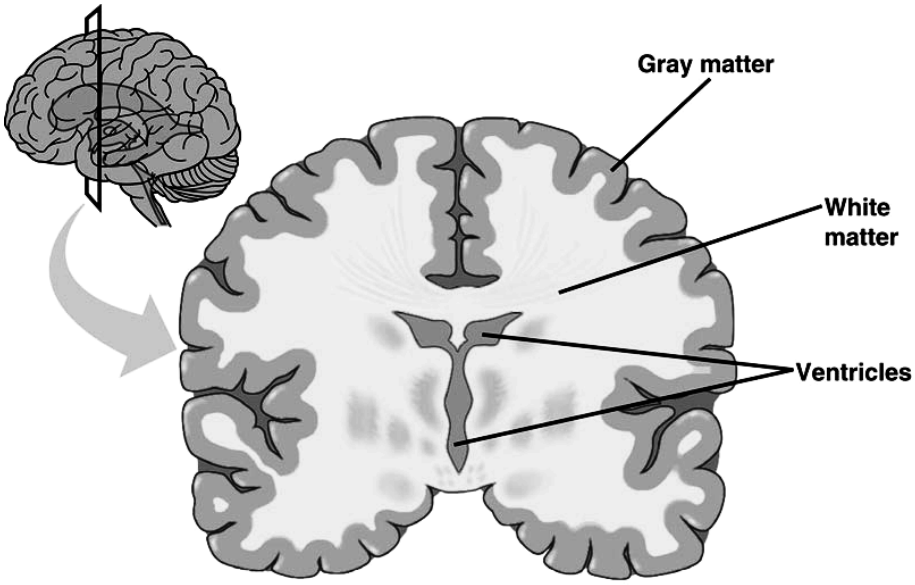

Figure 1.1: A schematic drawing of the human brain, indicating the two main tissue classes (gray matter and white matter) and the ventricles (filled with cerebrospinal fluid) on a coronal slice.

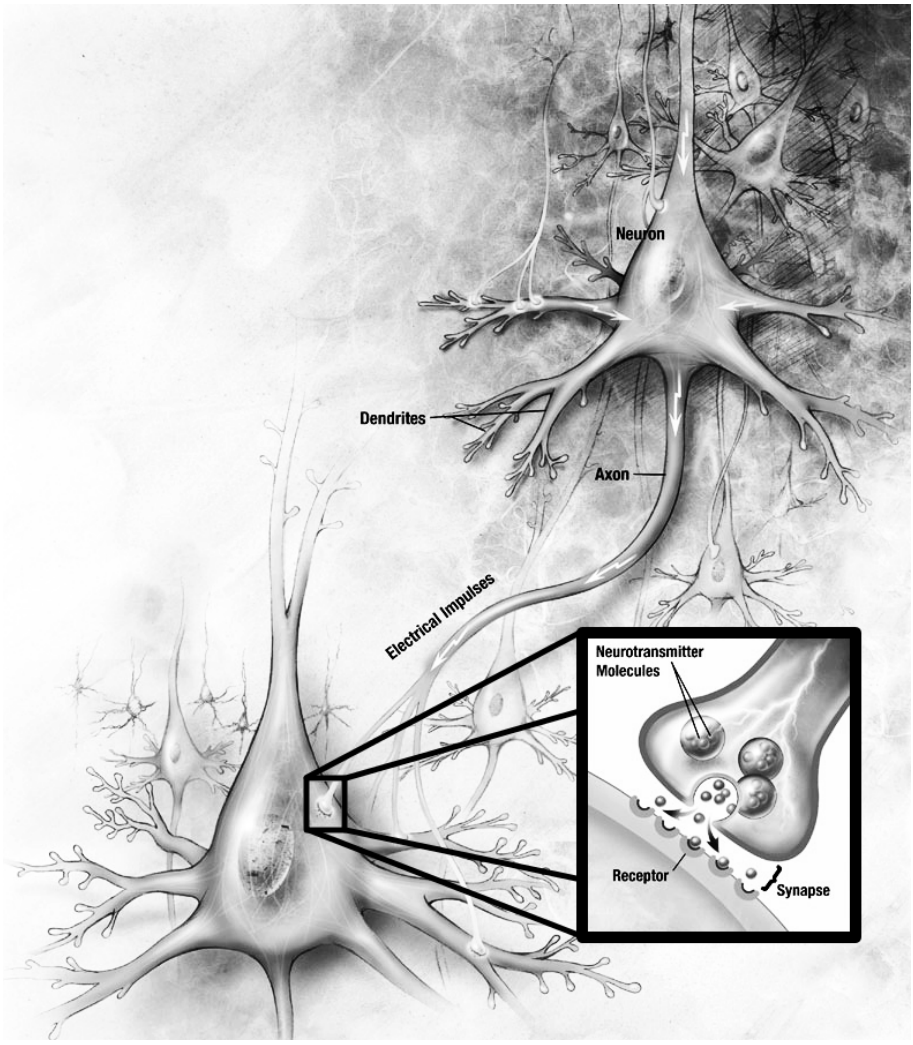

Figure 1.2: A stylized drawing of neurons. Note that the various features are not necessarily to scale: the axons can be much longer in comparison to the neuronal cell bodies.

The basic working unit of your brain is the neuron; a typical human brain contains a good 200 billion of them. Fig. 1.2 provides a general overview of the main parts of these neurons: the neuronal cell body, the incoming dendrites, and an outgoing axon. The features on the drawing are not necessarily to scale between each other: neuronal cell bodies ranging from $4\,\mu$m to $100\,\mu$m exist, while the axon diameter may vary from $0.2\,\mu$m to $20\,\mu$m, but can reach lengths from a few mm up to even 1 m. Information transfer along the axons themselves is an electrical process: electrical impulses travel at very high speeds along the length of the axons. Communication between neurons happens at the synapses. For most synapses, this is a chemical process: the axon terminal sends out neurotransmitters, and these subsequently bind to the receptors of the other neuron.

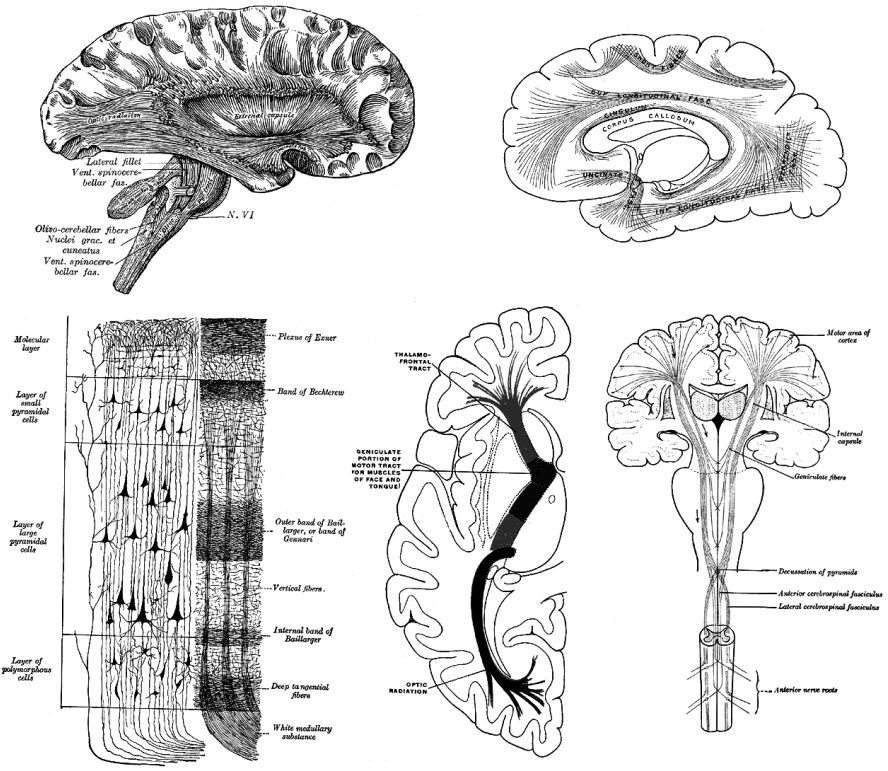

Figure 1.3: A collection of lithographs from "Gray's Anatomy", showing the well-organised structure of the white matter: big bundles of axons run in coherent patterns, connecting various regions of the gray matter.

While everything up to now might have sounded quite impressive already, the neatest fact is yet to come: the neurons are not just scattered around randomly within your brain; everything is actually quite nicely organized and, on a global scale, very similar among all normal healthy human beings! The gray matter mostly contains the bulk of the neuronal cell bodies and dendrites, while the white matter mostly contains (relatively) long axons. Also within the white matter itself, the axons are very well organized. They don't just run around randomly, but in several bigger and smaller coherent bundles. Fig. 1.3 presents a collection of lithographs from the famous "Gray's Anatomy" anatomy guide, which gives a very good impression of some of this large-scale axonal organization. While I introduced the brain earlier as being your central computing unit, a better comparison to modern technology would be the internet. All that white matter (Fig. 1.1) consists of very dense wiring, while the gray matter contains a massive amount of tiny computers. These bundles of axons are very dense too; if you'd cut one in half, observing a density of over 100000 axons per mm^2 is not exceptional. Even more stunning is the length of all this wiring: laying out your neurons one after the other will take you up to about 1000 km from your starting point. While this sounds intriguingly great (and greatly intriguing), the complexity and scale at which this all plays out is immense. How will we ever study such a huge amount of wiring, of which each single wire is only a few micrometers wide; and all of that is stored somewhere deep within our head? Surely the bundled and tightly packed nature of the axons might help us out. That, and diffusion.

1.2 Diffusion

So, what has diffusion to do with the brain then? Nothing in particular actually, but it might coincidentally help us to measure some of those complex structures we talked about before. For those who don't know yet: diffusion is something very simple in itself. Consider the experiment in the top row of Fig. 1.4. We start with a glass of clean water, and release a drop of ink somewhere in the middle of it. Everyone knows what to expect: over time the ink will slowly spread in the water. Wait for a (very) long time, and the whole glass' contents will even turn equally red. That's diffusion. Or, more specifically: that's the ink diffusing in the water. But why does it actually work like that? To understand this, we need to look at a smaller, molecular scale. A fluid such as water consists of molecules constantly moving around (even if the water is not flowing, such as the still water in the glass). With so many molecules moving around, collisions are unavoidable: molecules will collide and deflect each other's path. That's what's also happening to the ink's molecules: they are literally being pushed around by the water molecules. But of course, it also works the

other way around: the ink is also pushing around the water. And it goes even further: according to how we just defined diffusion, we might even say that some water molecules are diffusing among the other water molecules. That's self-diffusion. Just fill a glass with plain water, put it in front of you and look at it: self-diffusion is happening right in front of your eyes, but beyond a scale that your eyes can perceive of course. Because a molecule in the glass of water is equally free to move in any direction (*i.e.* it will not prefer any particular direction), we refer to this process as isotropic diffusion.

Now that we got that covered, we can scale up the complexity a bit. Consider the experiment in the bottom row of Fig. 1.4. This time, we've added some cylindrical structures in the water. Imagine that these are quite small and rather tightly packed together (*i.e.* the bottom row of Fig. 1.4 shows a large magnification of the real experiment). We again release a drop of ink, but somewhere among the cylinders. Molecules of both water and ink bump into each other just as before, but now they also hit the cylinders from time to time. It's quite straightforward to see that any diffusing molecule will have a

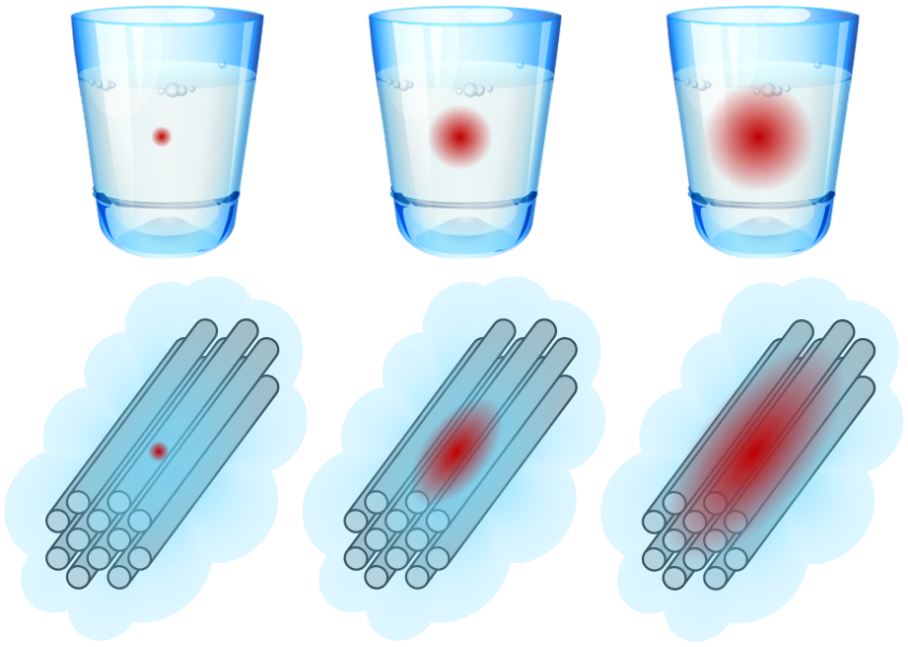

Figure 1.4: *Top row:* isotropic diffusion in a glass of water. *Bottom row:* anisotropic diffusion when cylindrical structures hinder diffusion.

harder time to move in directions perpendicular to the cylinders (because it has a high chance of hitting one) as compared to a direction along the cylinders. We say that diffusion is hindered, and anisotropic. The former refers to the cylinders (or anything else) just being there as a potential obstacle, while the latter refers to the fact that diffusion is not hindered equally in all directions. The result? The diffusing cloud of ink will feature an ellipsoidal shape, rather than a spherical one.

As the human body consists of a large amount of water, self-diffusion is also constantly happening in there. In the brain, however, we have this particular situation of dense bundles of axons being present. With some imagination, these are locally very similar to the cylindrical structures from our previous experiment. Thus, at a certain timescale, diffusion of water in between these axons will also feature some anisotropy, because the water molecules are hindered by the axons.

But the situation is even more pronounced! Let's take a look at Fig. 1.5. It presents a microscopic image of a slice through an axon. As you may notice, the axons appear to be tightly wrapped in some extra layers. These layers are

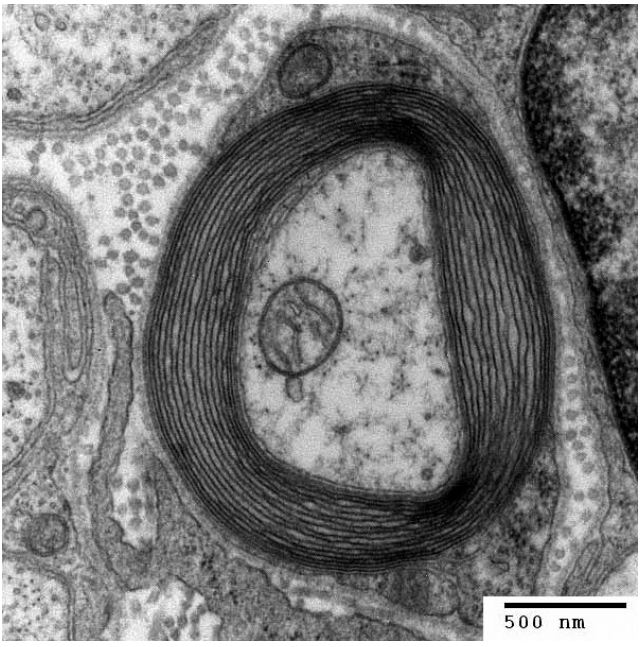

Figure 1.5: Transmission electron micrograph of a slice perpendicular to an axon, clearly showing the multitude of myelinated layers surrounding the axon.

actually the myelin (with high fat content) we mentioned before. Their specific function towards the axon is one of insulation. Similar to the plastic insulation we typically cover electrical wiring with, it will prevent the electric pulses from escaping the axon. Additionally, it will greatly increase the traveling speed of those pulses, which is essential for our brain to perform its complex tasks in an acceptable manner. So much extra layers of membranes might hinder diffusion even more. Finally, there is also water in the axon itself. While the water in between the axons had a hard time diffusing perpendicular to the axons because they hinder diffusion, the water inside the axons might even be fully restricted. Specifically for the water within an axon, an even more anisotropic pattern of diffusion may thus be expected.

The remaining question is: how will all of this help us? We established that tightly packed bundles of axons will cause anisotropic diffusion. The inverse reasoning is not necessarily valid, but let's still assume that, when we observe anisotropic diffusion within the brain, it is at least quite likely to be the consequence of axon bundles being present. So, if we were to have access to some probe to measure (the anisotropic nature of) diffusion, we might have a reasonable view on the microstructural axon bundles.

And now for the best part: there actually exists such a probe! It turns out that a magnetic resonance (MR) scanner is able to make measurements that are sensitive to diffusion in a limited timeframe. Combine this with the fact that we don't need to open up your skull to perform such MR measurements, and we can conclude that we might have access to a probe to (indirectly) measure white matter microstructure, *in vivo*. How this works, what data comes out, and how to deal with such complex information, is explained in the next chapter.

Chapter 2

From diffusion to the diffusion tensor

This chapter represents an adaption of a chapter I originally wrote as a contribution to a *practical* guide on diffusion tensor imaging (DTI):

- **Dhollander, T.**: From diffusion to the diffusion tensor. In: Van Hecke, W., Emsell, L., Sunaert, S. (eds.) Practical Handbook of DTI. To be published by Springer, Heidelberg.

Therefore, it bears a very specific style of writing, aimed towards a less technical audience. I never avoided the essential theory, however, but relied on intuition, examples and some repetition of the most important concepts to get the information across. A more informal style of writing was intentionally adopted for specific passages to balance out the unavoidable formal content, and as such keep the intended audience motivated to read on. On top of that, a great deal of text is *emphasized* – some parts even ***overemphasized*** – in order to stress new and important concepts, or other useful insights. While this should provide guidance for readers who are new to DTI, it can otherwise be safely ignored.

Intro

The term "diffusion tensor imaging" (DTI) is used on many occasions to informally refer to anything related to diffusion weighted imaging (DWI). However, this is already an excellent example of common careless, uninformed, and blatantly erroneous use of terminology; a source of confusion for newcomers to the domain. DTI is the practice of fitting a tensor model to the DWI data. It is one of the simplest ways to model the DWI data that accounts, up to some extent, for the anisotropy in this kind of data. Exploiting this anisotropy is key to obtaining the characteristic directionally-encoded color (DEC) maps and tractograms that are typically associated to the practice of DWI in general. Hence, it is not surprising many people use the term "DTI" in very different contexts. In this chapter, we aim to give the reader a feeling for what is really under the hood of the true art of DTI: obtaining these so-called diffusion tensors. What are they actually modeling? And, in this context, what *is* a tensor anyway? There's a short and clear answer to this: the diffusion tensor describes the apparent diffusion coefficient (ADC), in function of direction. Hmm... "ADC" you say...?

2.1 The (self-)diffusion coefficient

2.1.1 Measuring (self-)diffusion in the MR scanner

Sit down, relax and grab yourself a glass of water. Now put it in a nearby MR scanner (or rather, imagine doing this). Acquire a (non diffusion weighted) T2 weighted image as well as one of those fancy new diffusion weighted images. Now let's see if we can recover the diffusion coefficient of water from these 2 images. More accurately, we're talking about the *self-diffusion coefficient* here: it quantifies the freedom of movement of any single molecule of water, in the glass of water. We'll simply refer to it as D. Also note that we need (at least) 2 images: the diffusion weighted image would appear exactly the same as the T2 weighted image, if it were not *weighed down* by the appearance of diffusion; *i.e.* we're interested in the relative difference between both images. Since D should be the same in the entire glass of water, we simply choose one voxel. The intensity of the diffusion weighted image in this voxel will be referred to as S, while the (non diffusion weighted) T2 weighted image's intensity equals S_0. The process of diffusion should have caused *attenuation* in S, so S should always be smaller than S_0. The *decay* of S relative to S_0 is given by the so-called

Stejskal-Tanner equation (Stejskal and Tanner, 1965):

$$S = S_0 \cdot e^{-b \cdot D} \tag{2.1}$$

The b-factor in this equation captures all the relevant *scanning parameters* and was introduced to take abstraction of them (Le Bihan and Breton, 1985). In general, it can be seen as the amount of *diffusion weighting* that is applied; *i.e.* how sensitive the acquisition is to diffusion. Its value is typically set and reported in $\mathrm{s/mm^2}$. As a realistic value for our simple experiment at hand, we could have chosen *e.g.* $800\,\mathrm{s/mm^2}$. We can rewrite this equation so it becomes:

$$-\ln\left(\frac{S}{S_0}\right) = b \cdot D \tag{2.2}$$

The left side of the equation only contains measurements we obtained from the scanner (S and S_0), while the right side contains the scanning parameters in

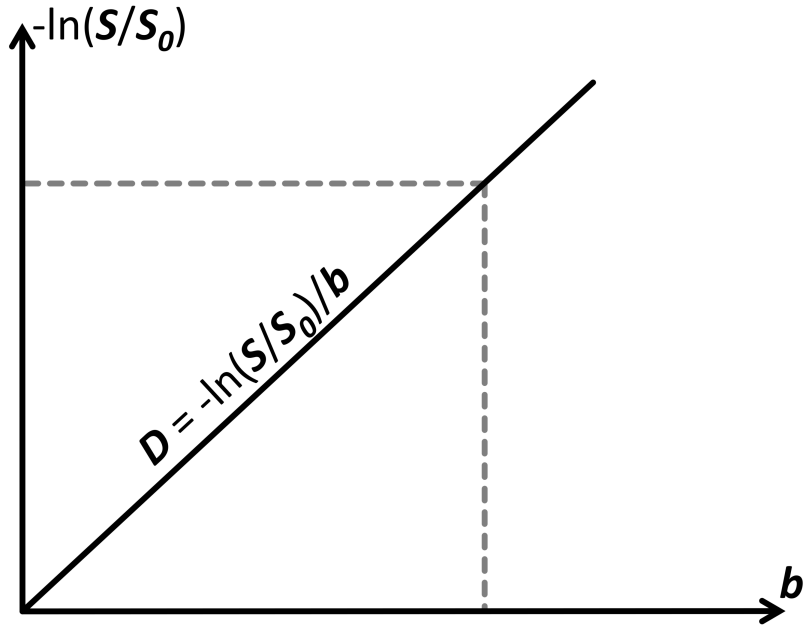

Figure 2.1: In case of free diffusion (*e.g.* in a glass of water, or CSF in the ventricles of the brain), the plot of $-\ln(S/S_0)$ in function of b-value is a straight line through the origin. One point (*grey short dashed lines*) is enough to fully fix this line. It requires 2 images (S and S_0) as well as knowledge of the b-value used to acquire S. The slope of the resulting line equals the self-diffusion coefficient D.

function of which we did so (all contained within the b-factor) and a constant (D for our glass of water at room temperature). From this we learn that a *logarithmic transform* (*i.e.* "$-\ln(\dots)$") of our *normalized measurement* (*i.e.* "S/S_0") depends linearly on the applied diffusion weighting (*i.e.* b). As they are both simply related by a factor D, plotting $-\ln(S/S_0)$ in function of b yields a straight line through the origin, as shown in Fig. 2.1. The slope of this line equals D:

$$D = \frac{-\ln\left(\frac{S}{S_0}\right)}{b} \tag{2.3}$$

The value of D is typically reported in mm^2/s. For our glass of water, $D = 2.2 \times 10^{-3}\,\mathrm{mm}^2/\mathrm{s}$ should be realistic at room temperature. If we were lying in the scanner ourselves and performed the above calculation for a voxel of cerebrospinal fluid (CSF) in the ventricles of our brain, a value of about $3.1 \times 10^{-3}\,\mathrm{mm}^2/\mathrm{s}$ is to be expected. While CSF consists mostly (99%) of water, the difference can be explained by our *body temperature*, which is of course higher than the normal room temperature. One point is enough to fully fix the slope of the line in Fig. 2.1 and per consequence also determine D. If we would have performed the measurement using a different b-value, we would obtain another point on this exact same line. Using a *higher b-value* would result in *more decay*, and thus a *lower value for S* (this can be most easily appreciated by looking at Eq. (2.1)). A *lower value for S* means a *higher value for* $-\ln(S/S_0)$. Consequently, we are simply considering a point further up the same line in Fig. 2.1.

2.1.2 Conclusions

We are now able to calculate the *self-diffusion coefficient* D of free water (be it in a glass or as CSF in the ventricles), using measurements from a MR scanner and the Stejskal-Tanner equation. The *minimum requirements* are a non diffusion weighted image (S_0), a diffusion weighted image (S) and knowledge of the b-value that was used for performing the acquisition of the diffusion weighted image.

2.2 The apparent diffusion coefficient (ADC)

2.2.1 Apparent complications

Feeling confident about the newly gained ability to obtain D from the 2 images we acquired of our brain, we also attempt to perform the same calculation in a

voxel of gray matter. Suddenly, however, we are confronted with a resulting value of about $0.9 \times 10^{-3}\,\mathrm{mm}^2/\mathrm{s}$. *Apparently*, the self-diffusion coefficient of water has changed, just because we measured it in the gray matter. Maybe something went wrong with the scan? We perform the acquisition again for a couple of different b-values. Carefully dotting out the obtained values of $-\ln(S/S_0)$ in function of b and connecting everything loosely by hand, we obtain a curve such as the one depicted in Fig. 2.2. *Apparently*, the self-diffusion coefficient of water now even changes in function of our chosen acquisition parameters. Using Eq. (2.3) to calculate D equates to connecting a certain measured point on this curve with the origin by a *straight line* (as shown in Fig. 2.2) and assuming its slope still equals D. Since the obtained values are clearly lower than expected and they also seem to vary in function of b, such a value is referred to as an *apparent diffusion coefficient (ADC)* (Le Bihan et al.,

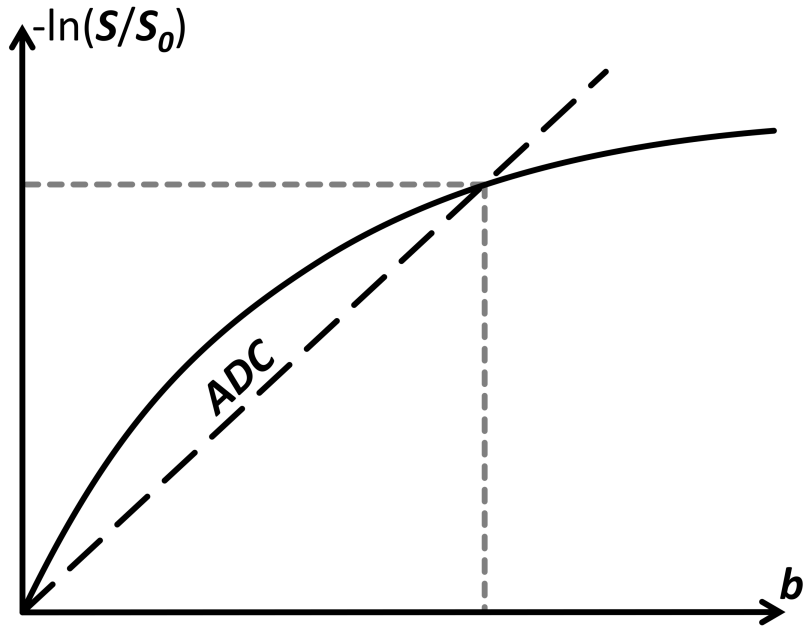

Figure 2.2: In case of hindered/restricted diffusion (in tissue, *e.g.* the grey matter of the brain), the plot of $-\ln(S/S_0)$ in function of b-value is a curve through the origin. Based on one point (*grey short dashed lines*), we can calculate an apparent diffusion coefficient (ADC). Just like D, it equals the slope of the line that connects this point to the origin (*black long dashed line*).

1986). It's calculated from the measurements in exactly the same way as D:

$$\mathrm{ADC} = \frac{-\ln\left(\frac{S}{S_0}\right)}{b} \tag{2.4}$$

To understand the behavior of the obtained ADC in regions containing tissue (*e.g.* gray matter), we need to look into how the acquisition of a diffusion weighted image works. An existing (*e.g.* T2 weighted) sequence is modified by adding a couple of diffusion sensitizing gradients. By taking abstraction of any complicated MR physics, we could say the MR scanner actually performs a simple experiment in each voxel: it takes a snapshot of all the water molecules, waits a bit, and then takes another snapshot. During the short waiting time, however, the molecules have the opportunity to *diffuse* a bit. Per consequence, a relative *displacement* of each molecule can take place in between both snapshots.

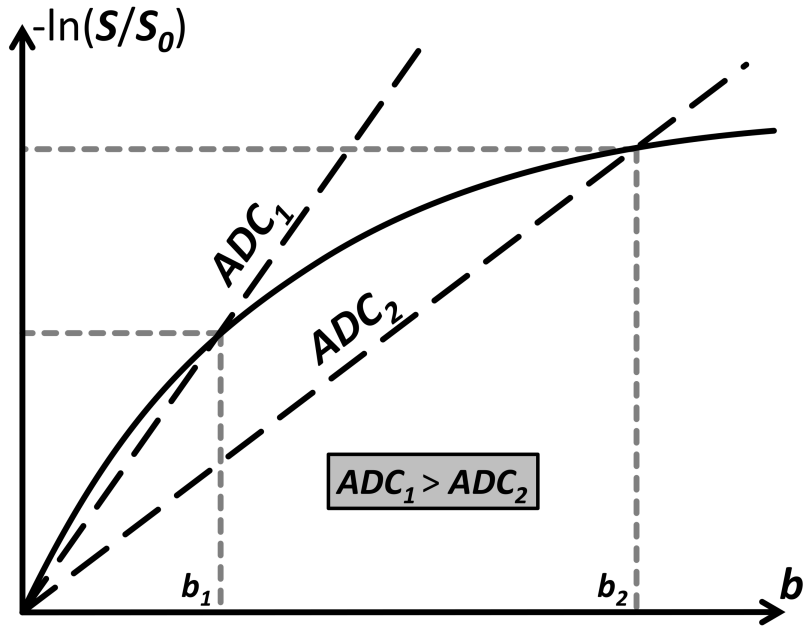

Figure 2.3: The ADC is dependent on the b-value used to acquire S. Due to the downwards curvature of the plot of $-\ln(S/S_0)$ in function of b-value, a larger b-value results in a lower ADC. An explanation lies *e.g.* in the fact that increasing the diffusion time allows more molecules to bump into cell membranes. This will on average decrease their final displacement, resulting in a reduced amount of attenuation of S and finally leading to a lower value for $-\ln(S/S_0)$ than expected in a free (non hindered/restricted) environment.

The expected signal of the original (*e.g.* T2 weighted) sequence is *attenuated* in function of the amount of *displacement* of all water molecules in the voxel (as well as the amount of applied diffusion weighting). From the measurements of such an experiment (relative to a non diffusion weighted image), the Stejskal-Tanner equation is able to reliably calculate D... if and only if nothing disturbs the experiment. However, in tissue, such as gray matter, there are *cell membranes* all over the place. Because the water molecules happen to bump into the cells – *i.e.* they are *hindered* – they have a harder time to diffuse further away during the experiment. Water inside the cells may even be *restricted* to a confined space. And thus, our calculation of D will *apparently* yield a lower outcome, which is why we call it the ADC instead. The time the experiment allows the molecules to diffuse, is one of the parameters that makes up the b-value. It's easy to imagine that a larger diffusion time will allow more molecules to hit some of these cell membranes. Hence, the effect of the *hindered/restricted* diffusion on our measurement will increase with b-value; yet another reason to refer to the outcome of our calculations using the term "ADC". This is also illustrated in Fig. 2.3: using a *larger b-value* renders the measurement of S *less sensitive to (truly) free diffusion*, in favor of hindered/restricted diffusion. As such, S will be *less attenuated* and the value of $-\ln(S/S_0)$ will be *smaller* than expected, yielding a *downward curvature* when plotting $-\ln(S/S_0)$ in function of b. This finally leads to an important property of the ADC in tissue, as indicated in Fig. 2.3: using a *higher b-value* results in a *lower ADC*.

2.2.2 Apparent advantages

At this point, you might start to wonder what the point is of trying to find out D in voxels containing tissue, only to end up having to deal with a deceiving ADC instead. However, you have to look at it from the bright side: we now effectively have access to a *probe* that tells us something about these *cells* that hinder/restrict diffusion. That's right: even though our voxel size might be quite crude ($2 \times 2 \times 2\,\mathrm{mm}^3$ or larger is not unusual), the measured values are sensitive to differences in structure at a *micrometer* scale! We are not interested in the ADC for the purpose of quantifying diffusion itself, but rather to *investigate properties of the tissue* that apparently caused the diffusion process to behave in the way that we measure.

Before moving on, let's investigate one more property of the ADC that teaches us something else about its capacity in distinguishing different tissues. Consider the setting in Fig. 2.4: it presents again $-\ln(S/S_0)$ in function of b, but this time for measurements at 2 different locations (*e.g.* in the brain). Looking at the plots and applying what we just learned, we can safely say that the voxel at position p_g contains more hindering/restricting tissue than the voxel at position

p_f. The former voxel's plot shows greater curvature, while the latter better approximates the straight line we would expect in case of free diffusion. Due to this difference in curvature of both plots, the *relative difference* in magnitude of $-\ln(S/S_0)$, and thus also ADC, *increases for larger b-values.* In other words, using a *larger b-value* results in a *better contrast* when calculating an ADC map. This fact of course begs the question why we should still limit ourselves to a certain b-value; *i.e.* why not use an absurdly high b-value for maximal contrast? The 2 most important factors that generally contribute to the b-value are the strength of the applied diffusion sensitizing gradients and the time that we allow the water to diffuse during the experiment. The former is limited by what we can achieve with available hardware. The latter is fully under our control. Allowing too long a diffusion time, however, might result in other more

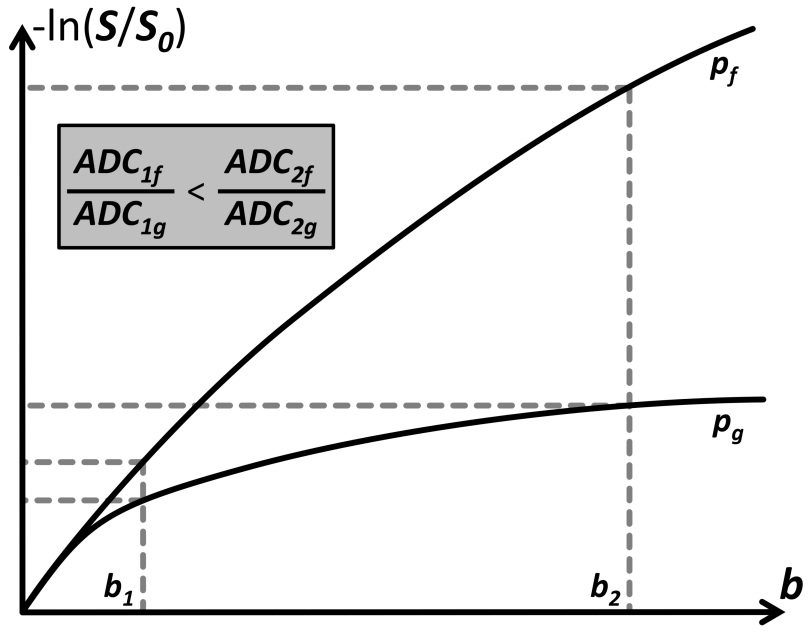

Figure 2.4: The contrast of the ADC, *e.g.* between 2 different tissues at voxel positions p_f and p_g, is dependent on the b-value used to acquire S. Due to different tissue properties, both plots of $-\ln(S/S_0)$ in function of b-value show a different curvature. The tissue at p_f imposes less hindrance/restriction on the diffusion as compared to the tissue at p_g and thus the accompanying curve is closer to a straight line. Hence, ADC_f is larger than ADC_g for a given b-value. Increasing the b-value also results in an increase of the relative difference between both ADC values, *i.e.* an increase of contrast.

macroscopic motion to be captured and thus confounding our measurements. Even if this were not the case, we also have to recall that S only *decays* further in function of b-value (remember Eq. (2.1) again?). The *noise level* of our measurements, on the other hand, *does not decrease*; *i.e.* using a *higher b-value* yields a *lower signal-to-noise ratio (SNR)*!

2.2.3 Conclusions

We have learned why the MR measurements in combination with the Stejskal-Tanner equation are not suited to calculate the true self-diffusion coefficient of water in voxels containing tissue, *e.g.* where diffusion is *hindered* or even *restricted*. The obtained *apparent diffusion coefficient (ADC)*, on the other hand, can provide interesting information about the microstructure of the tissue under investigation. The *minimum requirements* for obtaining it are again a non diffusion weighted image (S_0), a diffusion weighted image (S) and knowledge of the b-value that was used for performing the acquisition of the diffusion weighted image. The ADC is, however, dependent on the b-value: a *higher b-value* results in a *lower ADC*. It also *improves the contrast* (*e.g.* of the ADC map), but at the cost of a *reduced SNR*. Due to these dependencies, interpreting/reporting the ADC only makes sense when the b-value is also specified. Finally, comparing ADC values or maps originating from acquisitions with different b-values, does *not* make a lot of sense.

2.3 Gradient directions and anisotropy

2.3.1 Anisotropic complications

Up to now, we've been silently ignoring yet another important fact that will complicate everything even more. It concerns the diffusion sensitizing gradient: it's about time we started taking into account that it's applied along a certain *direction*. Nothing to worry about, if it were not for the fact that our measurements are only sensitive to diffusion along this direction. Actually, that is not fully correct; it's better to say that they are *only sensitive to diffusion with a component along this direction*. Before we start talking further about directions, let's settle on some reference frame. We define 3 (perpendicular) axes through the brain as follows: x runs from *left to right*, y from *back to front*, and z from *bottom to top*. So suppose we would apply the diffusion gradient along the direction of x, what are the implications then? It basically means that the measurements are *fully sensitive* to diffusion *along* x, but *gradually*

less sensitive to diffusion along directions that *increasingly deviate* from x, up to the point where they are *completely insensitive* to diffusion along directions *perpendicular* to x (*i.e.* directions in the yz-plane).

But why should we worry about directionality of diffusion anyway? In our earliest experiments with a glass of water or CSF in the ventricles, we shouldn't: diffusion takes place *equally in all directions*. In tissue randomly containing cells – imagine a bunch of spherical cells packed together – diffusion is hindered and restricted, yet probably more or less *equally in all directions*. So again, there's nothing to worry about: as we are in both cases studying *isotropic* measurements, it is sufficient to only measure along a single direction. Our findings (*e.g.* calculating the ADC) should have been the same for measurements along any other direction. But let's consider the more interesting case of the white matter in the brain: it consists of long coherent bundles of axons, almost resembling a bunch of cylindrical tubes packed closely together. One can imagine that water molecules in between and inside these tubes have *an easier time diffusing along them rather than perpendicular to them*. We thus say that diffusion in the white matter is *anisotropic*.

But how relevant is this? Is this anisotropy large enough to be measured; *i.e.* can we see it in our diffusion weighted images? To answer this question, we'll introduce some real data. In Fig. 2.5, we start by presenting a classic T1 weighted and T2 weighted image for reference. Next is the non diffusion weighted image: it's again a T2 weighted image, but it already shows the *lower spatial resolution* at which DWI datasets are typically acquired. In this particular dataset, the voxel size equals $2.2 \times 2 \times 2\,\mathrm{mm}^3$. Because the image is not diffusion weighted, but it is acquired as part of a DWI dataset, we also sometimes (informally) refer to it as the "B0" (it equals a diffusion weighted image with a b-value of 0). For convenience, we already applied a whole brain mask to it. On the second row of Fig. 2.5, 3 diffusion weighted images (DWIs) are shown (also masked). They were all acquired using exactly the *same amount* of diffusion weighting: $800\,\mathrm{s/mm}^2$. The diffusion sensitizing gradients are, however, applied *along different directions*: respectively along the direction of x, y and z. Differences in contrast can clearly be seen, which consequently confirms that we will have to account for *anisotropy* in our measurements.

2.3.2 Anisotropic advantages

Just as when we introduced the ADC, we'll also try to use this fact to our advantage: we now have access to a probe that might even provide us with information on the *anisotropy of the microstructure* that hinders and restricts the process of diffusion. Applying what we have learned from this chapter up

to this point, let's see if we can already figure out something useful from these 3 diffusion weighted images. Consider the indicated region in the genu of the corpus callosum (GCC): it has *a low DWI intensity along* x, but a (relatively) *higher DWI intensity along* y *and* z. Because we know that more diffusion causes increased decay of S (the DWI intensity), we can conclude from these images that there is *more free diffusion along* x, while there is *more hindrance and restriction along* y *and* z. Translating this to "reality", we might infer that *a bundle of tube like axons runs along the left-right axis* in this region, connecting both hemispheres of the brain. Note that we are applying *inductive* reasoning here: we know that such a left-right oriented structure would result in such a pattern of diffusion and thus also such DWI measurements along these 3 directions, yet we reason that the latter measurements were effectively caused by the former structure. Considering we only measured along 3 directions, that's

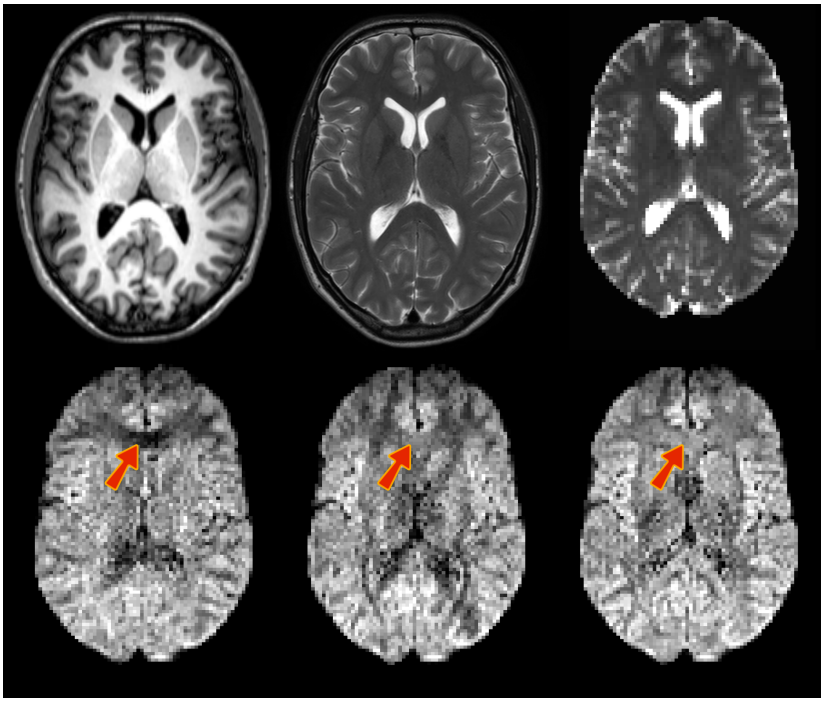

Figure 2.5: *Top row*: T1 weighted image, T2 weighted image, B0 image ("diffusion weighted image" with a b-value of 0, *i.e.* non diffusion weighted). *Bottom row*: diffusion weighted images (DWIs) acquired by applying gradients along the direction of x, y and z. The arrows indicate a region in the genu of the corpus callosum (GCC), where the anisotropy can be easily seen and understood.

a pretty strong conclusion. Of course, inherently we might have also applied possible anatomical knowledge and the fact that a structure along this direction makes sense considering the spatial/anatomical neighborhood of the region (*i.e.* the region is in between both hemispheres).

2.3.3 Getting a grip on the information overload

In practice, however, we will typically perform the acquisition using more than 3 different gradient directions. In the previous example, we were just lucky that the structure under investigation accidently happened to run along one of the 3 mutually perpendicular directions that we sampled. If it would instead be running at any other oblique angle, these 3 measurements would clearly be *inadequate* to determine its direction. In our experiment at hand, however, we actually acquired DWIs for a total of *45 different gradient directions*! The specifics of such an acquisition are presented in the *gradient table*, that contains *one row for each acquired image*, representing its gradient direction and b-value. The gradient table for our current experiment is provided in Fig. 2.6. As it was already quite tedious to infer information by mentally combining 3 images, considering 45 DWIs all at once is nearly impossible. To begin with, we will no longer visualize the original DWIs, as they are still only representing a (partially) decayed T2 weighted signal. Because of this, these DWIs suffer so-called *T2 shine-through*: a higher intensity might not (only) result due to hindered/restricted diffusion; it might also be caused by an originally high T2 intensity (*e.g.* in areas containing CSF). Therefore, it is more evident to consider DWIs after *normalization by the B0* (*i.e.* the normalized measurements "S/S_0").

Such normalized versions of our original DWIs for the 3 x, y and z gradient directions are shown in Fig. 2.7. Next up is the actual challenge of visualizing the information of all 45 normalized DWIs in a conveniently organized way. Rather than showing 45 separate images, we could try to combine all information of each single voxel and visualize it within that particular voxel. Since the different values of S/S_0 are a *function of the gradient direction*, a *spherical polar plot* is the perfect candidate for the job. In such a plot, the radius of a sphere is locally manipulated to equal the function value at that 3 dimensional angle. We also smoothly interpolated the values between the 45 directions in order to achieve the final visualization in Fig. 2.7. Note that, due to the multitude of information on display, we have to zoom in up to a reasonable level to show everything with the required amount of detail. We choose to further focus on the region of the GCC that was the subject of our earlier thought experiment. Furthermore, a little extra *color* was added to the plot: each point on the surface of the spherical polar plots is colored according to its *direction*: *red* is assigned to x, *green* to y and *blue* to z. In (the middle of) the GCC, we spot larger values

0	0	0	0
0.2094047171315730	-0.9767227295418320	-0.0465013337378605	800
-0.4750314453755630	-0.6955464694333420	-0.5390363946555150	800
0.6664846048909360	-0.0445988907177365	0.7441835864826100	800
-0.3582039903915180	0.6662077459922660	-0.6541078966405300	800
0.2677922332193600	-0.0027000724215939	-0.9634729002085670	800
-0.9857919430003580	0.1144991316171740	-0.1228991211298600	800
0.6177990677639750	0.3739997525290050	0.6916997159014940	800
-0.6308643964634750	0.4140767555916230	-0.6561635114462630	800
0.6205685840497420	0.1912902907213110	-0.7604621339463620	800
-0.3779719889511070	-0.1126918360486610	-0.9189329277245030	800
0.3395100689024380	-0.2920086911190300	0.8941274167728070	800
-0.7758499588418720	0.3188207857101360	0.5444356233424090	800
0.5160887818975360	0.5839874957334980	-0.6265867649627110	800
-0.7296728779714800	0.6215772630201830	0.2849898195535330	800
0.8361102177502960	0.5423069132347100	-0.0826009420732837	800
-0.2423039673649470	-0.8370139550108740	0.4906082210035440	800
0.6098996127834700	0.7761999371383710	0.1598002500403600	800
-0.2550936809864850	-0.9252774535964550	-0.2806935121922570	800
0.1169006611114160	0.7813049357224980	0.6131042593616330	800
-0.5508125092269130	0.8346194207164320	0.0040002802549276	800
0.8721102831781370	-0.2669032939932760	0.4101052128789500	800
-0.7807080188188620	-0.4384045893749320	0.4453048454339160	800
0.0925995394001606	0.6432967345937930	-0.7599964714153890	800
-0.5103136738555730	0.5626154590887860	0.6504181727706740	800
0.0168993374063531	-0.0846967117840156	0.9962634588341590	800
-0.8512463460131440	0.5210674735188420	-0.0621960323490279	800
0.2549975421500550	0.4799957213668850	0.8393928525826960	800
-0.8612898527581150	-0.1205986881754780	-0.4935947183127650	800
0.2075896523172570	-0.4157796321621520	-0.8854568502924950	800
-0.9434712867959140	0.2488926025139570	0.2188935892288590	800
0.3534894353454850	-0.8671746164237320	0.3507896858938190	800
-0.0499972977032263	0.5097729847185540	-0.8588549203873210	800
0.6819782352946690	-0.6446797384480750	0.3453892317645280	800
-0.5366899164675200	0.1913966296310320	0.8217854122141420	800
0.3007924095658790	0.9508764278970280	-0.0731979864295444	800
0.0183995925409323	0.9656769222860070	-0.2590936833628960	800
0.7972628913636550	0.4963771701857660	0.3434844785035220	800
-0.0140999738954793	0.9389002412956230	0.3439004618071470	800
0.3105054355472360	0.3664065037097980	-0.8771160975234600	800
-0.9615751036584600	-0.1524960776055520	0.2282942538460920	800
0.9634198699916120	0.2308048753520390	0.1362030235313290	800
-0.1174026699519580	0.7255170198882160	0.6781162635866160	800
0.5724948174210340	0.7776932591366530	-0.2596976679130320	800
-0.8347242096433000	-0.0223005675802262	0.5502164833235800	800
0.3981243181229750	-0.8272510935920460	-0.3964248421452410	800

Figure 2.6: The gradient table contains one row for each acquired image. The x, y and z components of the gradient direction are provided in the first 3 columns, while the b-value is given in the last one. A b-value of 0 indicates a B0 image; the gradient direction is irrelevant in such a case. The *red encircled rows* refer to the DWIs presented in Fig. 2.5.

for green (y) and blue (z), and smaller values for red (x). By linking *larger values* to *hindrance/restriction*, we can thus confirm our hypothesis of an axonal bundle connecting left and right.

Associating larger values with less diffusion still feels a bit awkward, to say the least. So why don't we simply employ the ADC values? Easy enough: just calculate the 45 ADC maps from the normalized DWIs using Eq. (2.4). We present these maps – again for the 3 x, y and z gradient directions – in Fig. 2.8. This time, *larger values* equal *more free diffusion*. We can just as well create a spherical polar plot of the 45 ADC values in each voxel, which is again provided for the GCC region in Fig. 2.8. *Larger values* are now conveniently oriented along the direction of the *greatest amount of free diffusion*, and colored accordingly. Finally, remember that property of the contrast increasing with

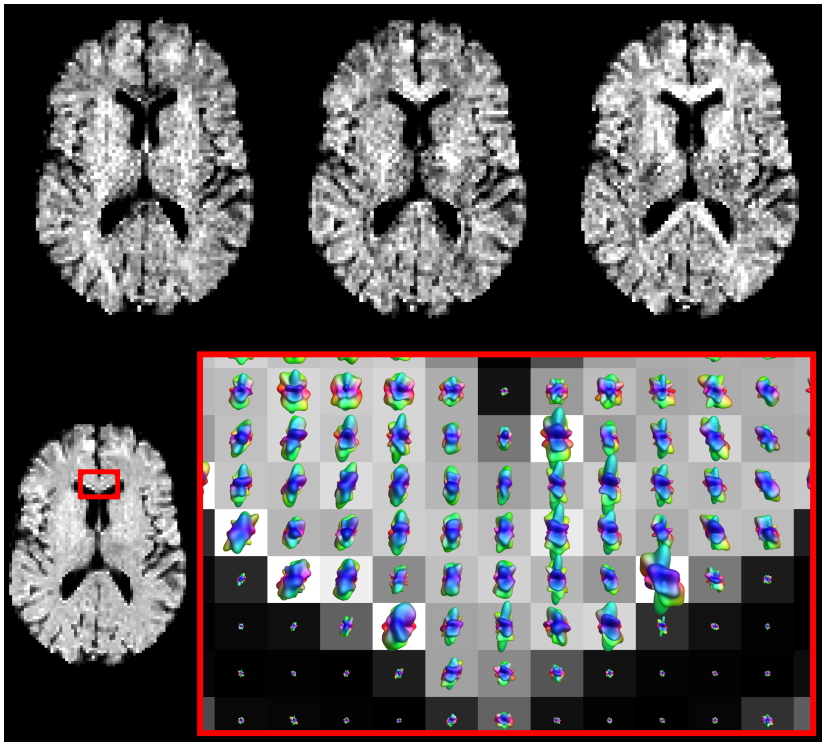

Figure 2.7: *Top row*: DWIs for the x, y and z gradient directions, after normalization by the B0 (*i.e.* the normalized measurements "S/S_0"). *Bottom row*: spherical polar plots of the normalized DWI values in a region of the GCC, overlaid on a map of the average normalized DWI value.

b-value? Of course, it also applies for measurements (and ADC values) acquired using different gradient directions: using a *higher b-value* will *increase the contrast* in these spherical polar plots. However, as we reasoned before, the SNR will also drop.

2.3.4 Conclusions

We started taking into account the fact that the diffusion sensitizing gradient is applied along a certain *direction*. The resulting DWI measurement is only sensitive to *diffusion with a component along this direction*. From DWIs acquired using different gradient directions, we could conclude that *anisotropic diffusion* takes place in the *white matter* up to a measurable extent. Again using this to our advantage, we now have a probe for the *anisotropy of microstructure* in

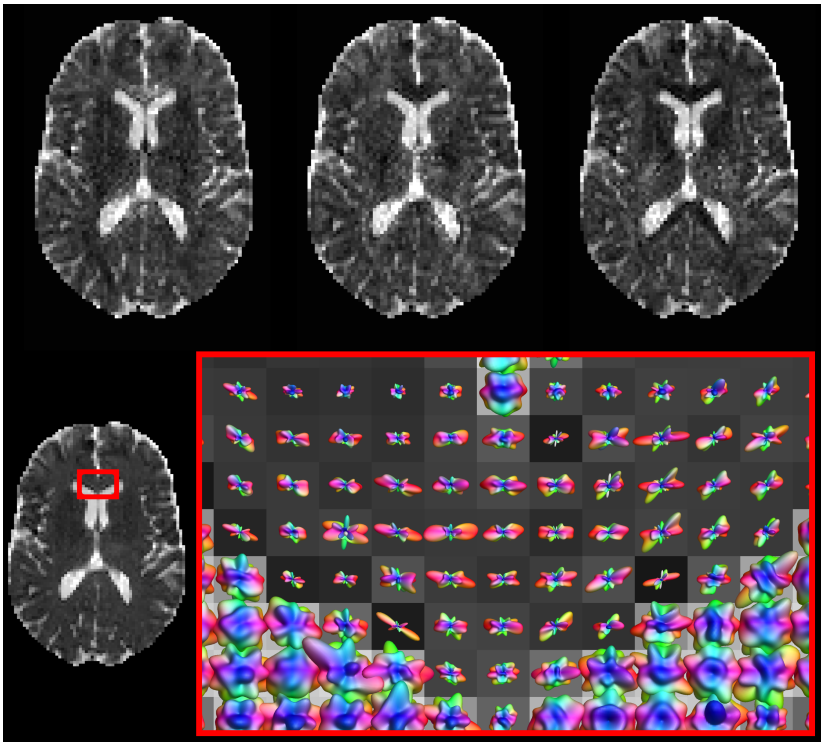

Figure 2.8: *Top row*: ADC maps for the x, y and z gradient directions. *Bottom row*: spherical polar plots of the ADC values in a region of the GCC, overlaid on a map of the average ADC value.

each voxel. There are different ways to visualize data resulting from acquisitions using many different gradient directions, yet the most convenient option was *a spherical polar plot of the ADC values* in each voxel: such a visualization shows *larger values* along the direction exhibiting the *greatest amount of free diffusion*. Optional *color coding* is typically done according to a directional scheme: *red for x (left-right), green for y (back-front) and blue for z (bottom-top)*. The *requirements* for investigating the anisotropic nature of diffusion are a B0 (non diffusion weighted image), a number of DWIs and knowledge of the *b-value and* ***gradient directions*** that were used for performing the acquisition of the DWIs. ***This latter point is very, very important!*** Did we just stress that enough? Because it is (very, very important): without the accompanying *b-value and* ***gradient directions***, the full set of carefully acquired DWIs is nigh *useless*; we wouldn't be able to associate the (normalized) measurements nor the ADC values with any directions. This vital piece of information should thus be stored with the data; it is often summarized in a gradient table, as shown in Fig. 2.6. On the number of gradient directions: more measurements are of course always better, yet require more scanning time. And finally, a *higher b-value* yields *better contrast* – also in the spherical polar plots of *e.g.* the ADC values – but will *reduce SNR*.

2.4 The (apparent) diffusion tensor

2.4.1 Motivation and implications of modeling

Looking back at the spherical polar plots of the ADC values in Fig. 2.8, we notice that they appear quite noisy. That's not surprising, since they simply present a logarithmic transformation of the original (normalized) data: nothing is modeled, all the measurement noise is still showing (albeit logarithmically transformed... remember this, as it will happen to bug us later on). And thus models were invented. Without going into the how and why of some historical choices that have been made in model development, we'll just introduce the (legendary) *diffusion tensor model* (Basser et al., 1994a) that is central to the theory and practice of DTI. In this context, to be exact, we should refer to it as the *apparent* diffusion tensor. This name refers to the fact that we will employ a tensor to represent/model *the values of the ADC* in function of (gradient) direction, in each voxel. This means that, once we have somehow determined the correct parameters of this model in each voxel, we can evaluate it for as many directions as we like in order to visualize it again as a spherical polar plot of *(modeled) ADC values*. As the diffusion tensor model has *only 6 parameters*

(compare this to the 45 ADC values we just obtained from our dataset in each voxel!), it will greatly simplify the features of our directional profile of the ADC.

2.4.2 Understanding DTI, in theory: the maths!

Mathematics... it's not as hard as it sounds, so let's just get to it then! From this point on, we will represent a (gradient) direction by a 3 element column *vector* $\mathbf{g}$, and the apparent diffusion tensor $\mathbf{D}$ by a 3×3 *symmetric matrix*:

$$\mathbf{g} = \begin{bmatrix} g_x \\ g_y \\ g_z \end{bmatrix} \qquad \mathbf{D} = \begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{xy} & D_{yy} & D_{yz} \\ D_{xz} & D_{yz} & D_{zz} \end{bmatrix} \tag{2.5}$$

For the mathematics (and software that employs it) to work out well, $\mathbf{g}$ should be a *unit vector*. As stated before, the tensor $\mathbf{D}$ has *only 6 free parameters* (the tensor elements D_{xx}, D_{yy}, D_{zz}, D_{xy}, D_{xz}, D_{yz}) because its matrix is *symmetric*: the elements above and below the main diagonal are the same. We'll get into the meaning of these separate tensor elements later. Given such a tensor $\mathbf{D}$, we can "*evaluate*" it for a given direction $\mathbf{g}$ by using the following expression:

$$\mathbf{g}^\mathsf{T}\mathbf{D}\mathbf{g} = g_x^2 D_{xx} + g_y^2 D_{yy} + g_z^2 D_{zz} + 2g_x g_y D_{xy} + 2g_x g_z D_{xz} + 2g_y g_z D_{yz} \tag{2.6}$$

where $\mathbf{g}^\mathsf{T}$ is the transpose of $\mathbf{g}$. The right side of the equation simply shows what you would obtain if you did the symbolic math by hand using the vector and tensor element symbols from Eq. (2.5). The outcome of this expression – if we were to fill in some specific numbers representing the vector and tensor elements – is thus a single scalar number: *the value of our tensor model, along a given direction*. As we will now employ such a tensor to symbolize the ADC values, we can simply plug it into the good old Stejskal-Tanner Eq. (2.1) to obtain the following expression:

$$S = S_0 \cdot e^{-b \cdot \mathbf{g}^\mathsf{T}\mathbf{D}\mathbf{g}} \tag{2.7}$$

Don't take this one lightly: **this is the *essence* of DTI**. It provides the direct relationship between the chosen experimental parameters (b and $\mathbf{g}$), the measurements (S and S_0), and the parameters of the diffusion tensor model ($\mathbf{D}$). It now effectively includes the gradient direction $\mathbf{g}$ that we took abstraction of before, while the vehicle to describe the ADC is no longer a single number, but a *tensor* that can describe values that vary in function of (gradient) direction. Just like we did before with the Stejskal-Tanner equation, we can rewrite Eq. (2.7) to single out the parts that equate to the ADC:

$$\mathbf{g}^\mathsf{T}\mathbf{D}\mathbf{g} = \frac{-\ln\left(\frac{S}{S_0}\right)}{b} \tag{2.8}$$

The right hand side equals the expression of the ADC that we introduced before (*i.e.* Eq. (2.4)), while the left hand side simply says that we would like to see this ADC value arising from our model $\mathbf{D}$ when evaluated for the gradient direction $\mathbf{g}$ that this particular ADC value relates to. Completely writing out the left hand side expression using Eq. (2.6) finally yields the following result:

$$g_x^2 D_{xx} + g_y^2 D_{yy} + g_z^2 D_{zz} + 2g_x g_y D_{xy} + 2g_x g_z D_{xz} + 2g_y g_z D_{yz} = \frac{-\ln\left(\frac{S}{S_0}\right)}{b} \quad (2.9)$$

If we now perform a DWI experiment as before (acquiring S and S_0, carefully noting down b and $\mathbf{g}$), we can fill in everything but the *6 unknown parameters* of the diffusion tensor model. As solving a single equation for 6 unknowns is quite an impossible task, we'll clearly need more of these equations, and by consequence more acquisitions. Mathematically, we know that at least 6 equations will be necessary to determine the full apparent diffusion tensor. In practice, we'll be needing *at least* 6 DWIs for *different* gradient directions as well as a single B0 to normalize our measurements to. As stated before, it is essential that every DWI is tied to its respective gradient direction. Acquiring *more* DWIs (for different gradient directions) will lead to more than 6 equations. In such a setting, no exact solution for the 6 unknown tensor elements generally exists, because the full system of equations is *overdetermined*. This actually is a good thing! Even though we can find a single exact solution in case of 6 DWI measurements, this solution will also exactly represent all the *noise* in the data. If we perform the acquisition using a *larger* amount of different gradient directions, a solution will have to be found that *fits* the data as well as possible. We then hope that the part that *doesn't fit* the model (*i.e.* the *residuals*) is the *noise*, which we (optimally) don't want to model anyway. We'll focus on the issue of tensor fitting later. For the time being, let's take it for granted.

2.4.3 Understanding the tensor elements, in practice

So, effectively applying such a tensor fitting method to our 45 gradient direction dataset at hand, we end up with 6 numbers in each voxel, *i.e.* the components which describe an apparent diffusion tensor. Now what can we actually do with it? For starters, let's take a look at some maps of these 6 tensor elements. There are 2 distinct categories amongst them: the *diagonal* elements (D_{xx}, D_{yy}, D_{zz}) and the *off-diagonal* elements (D_{xy}, D_{xz}, D_{yz}). The former are presented on the first row of Fig. 2.9, while the latter are shown on the second row. The interpretation of the *diagonal* elements is straightforward: they represent ADC values along the respective directions of x, y and z. Because we've shown maps of the original (unfitted) ADC values along these directions in Fig. 2.8, we can compare them directly to the maps of D_{xx}, D_{yy} and D_{zz} in Fig. 2.9. Indeed,

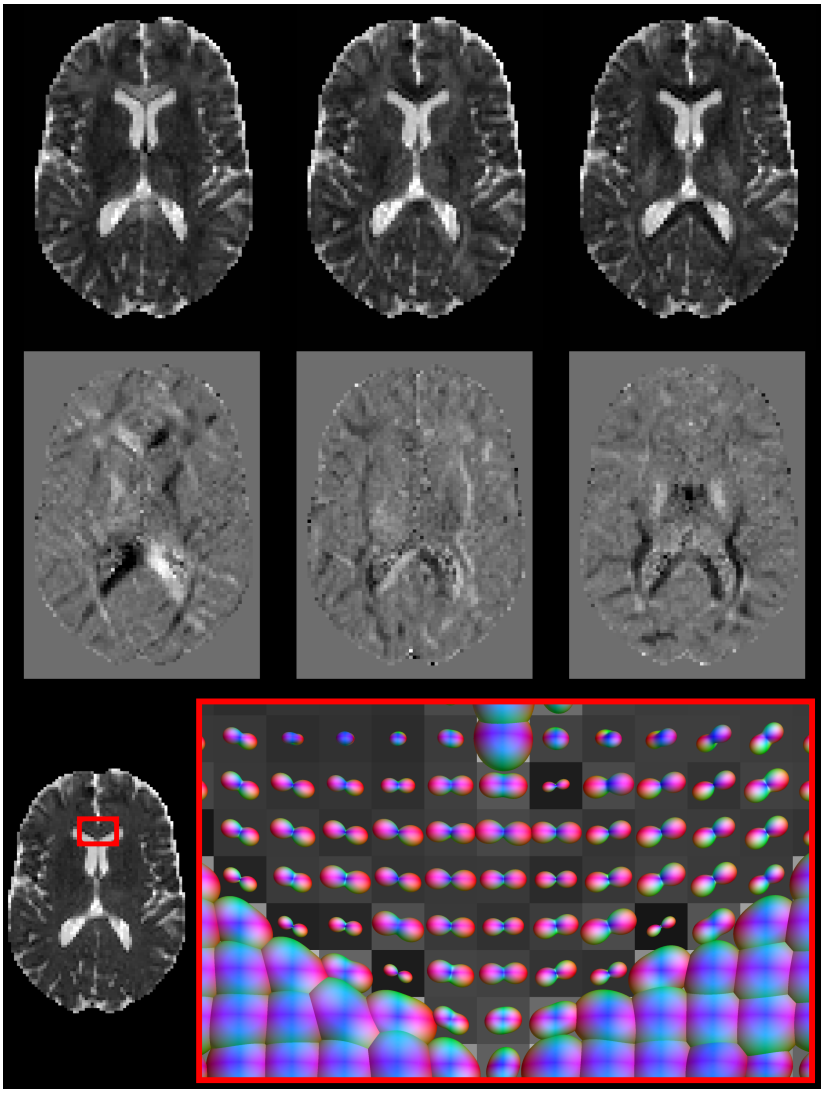

Figure 2.9: *Top row*: maps of the diagonal diffusion tensor elements (D_{xx}, D_{yy}, D_{zz}). *Middle row*: maps of the off diagonal diffusion tensor elements (D_{xy}, D_{xz}, D_{yz}). The background grey level equals zero; darker/brighter levels represent negative/positive values. *Bottom row*: spherical polar plots of the ADC values provided by the diffusion tensor model in a region of the GCC, overlaid on a map of the average ADC value. Note the characteristic peanut shapes that appear in the GCC.

they look more or less alike. The more careful observer may note that the latter look less noisy. This is not surprising: they represent fitted values, *i.e.* all 45 measurements contributed to them. From what we already know, we can even figure out mathematically why *e.g.* D_{yy} corresponds to the value of the tensor model (*i.e.* the ADC) along y: just fill in $[0\ 1\ 0]^\mathsf{T}$ (*i.e.* the direction of y) as direction $\mathbf{g}$ in Eq. (2.6) and evaluate using the right hand side expression; the outcome trivially equals D_{yy}. We can thus conclude that the *diagonal* elements are in practice meaningful and quite easy to understand. The *off-diagonal* elements (second row of Fig. 2.9), on the other hand, offer a less intuitive source of information. They represent the covariance between each pair of axes (*i.e.* xy, xz and yz). That's because the full diffusion tensor is actually a covariance matrix. Apart from being a great conversation starter at an engineering party, those last 2 sentences won't get you anywhere in daily practice: the *off-diagonal* elements just don't really have a *direct practical use* or meaning. The only reason we do discuss them here, is to emphasize what they *don't mean*: they do *not* represent the values of the ADC along some diagonal direction (so don't mistake them for that!). This should also be clear from the fact that they equally cover a range of positive as well as *negative* values (while ADC values should *not* be negative). One final property worth mentioning though: if all off-diagonal elements are zero, the tensor is perfectly *aligned* to the x, y and z axes. What that means, will become more clear if we visualize the tensor in 3D.

2.4.4 Understanding the tensor, in practice: peanuts!

Talking about visualization, let's take a look at a spherical polar plot of the ADC values actually represented by the fitted diffusion tensors. These are shown for our trustworthy GCC region on the bottom row of Fig. 2.9, and can again be directly compared to the original (unfitted) values in Fig. 2.8. From this comparison, it is obvious that the noisy appearance has been greatly reduced: while the original plots showed a unique and different pattern in each voxel (due to the varying noise), the directional profiles that represent the tensor fitted values are much more *consistent* within regions. In regions where a single bundle of axons is present, *e.g.* the GCC, these plots typically take on the shape of *peanuts*. The advantage of modeling is that some features of interest, such as the *main direction* of the tensor, are recovered more prominently. From the region shown in Fig. 2.9, it is now also more evident that the nearby CSF in the ventricles exhibits an isotropic pattern of diffusion. And finally, for those who'd like to go just that extra mile in interpretation: consider again one of those curious maps of the *off-diagonal* elements, D_{xy}, and note that it indeed shows a value of *zero* for the voxels right in the middle of the GCC, where the main directions of the tensors are nicely *aligned* to the x axis.

2.4.5 Conclusions

We have introduced the (apparent) *diffusion tensor* model (Basser et al., 1994a), which is used in DTI to represent the *ADC values* of our measurements along different (gradient) directions. The diffusion tensor $\mathbf{D}$ is represented by a 3×3 *symmetric matrix*, containing *6 unique tensor elements*, and can be evaluated along any direction $\mathbf{g}$ by the expression $\mathbf{g}^{\mathsf{T}}\mathbf{D}\mathbf{g}$. Casting this expression in the role of the ADC value in the Stejskal-Tanner equation, yields the one and only equation at **the *core* of DTI**: it directly relates the experimental parameters, the measurements and the parameters of the diffusion tensor model to each other. This equation can again be rewritten to show clearly that the *diffusion tensor* model is meant to fit the *ADC values*. As there are now *6 unknowns* in this equation, the *minimum requirements* for obtaining the diffusion tensor are a B0 (non diffusion weighted image), at least 6 DWIs and knowledge of the *b-value and* ***gradient directions*** that were used for performing the acquisition of the DWIs. We simply cannot stress enough that knowledge of the *b-value and* ***gradient directions***, *i.e.* the full gradient table as shown in Fig. 2.6, is absolutely *essential* to fill in the equations and obtain the diffusion tensors! Fitting the tensor model to data with a larger (than 6) number of DWIs *reduces the noisy appearance of the ADC values* when visualized as a spherical polar plot. In regions of white matter containing a single consistent bundle of axons, the plot has a characteristic *peanut* shape that clearly shows features such as the *main direction* of the tensor. The *diagonal* elements of the diffusion tensor represent *ADC values along the x, y and z axes*; while the *off-diagonal* elements represent the covariance between pairs of those axes. Maps of the former thus provide a *meaningful* interpretation, while maps of the latter are neither intuitive nor useful in daily practice. Just *don't mistake the off-diagonal elements* for ADC values along some oblique angle. A final overview of the most important steps taken up to this point is shown for a single voxel in the middle of the GCC in Fig. 2.10: from raw DWI measurements, to calculated ADC values, and finally the fitted tensor!

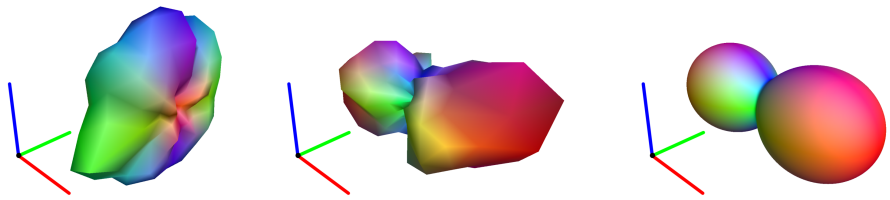

Figure 2.10: From DWI data to the tensor, for a single voxel in the middle of the GCC. Spherical polar plots of the DWI values (*left*), the ADC values (*middle*) and the ADC values evaluated from the fitted tensor (*right*).

2.5 Eigenvalues and eigenvectors

2.5.1 The tensor elements: not very practical

Looking back at the ADC peanuts that represent the diffusion tensor values in Fig. 2.9 and Fig. 2.10, we notice that the tensor model indeed did a good job in capturing the most important *features* of the angular ADC profile: we can clearly observe its main direction, the maximal ADC value (along this main direction), etc... and how these features relate to each other over larger regions (*e.g.* qualitatively observe the curving global path of the axon bundle in the GCC). Although these very practical features are in each voxel captured and described by those 6 unique tensor elements, it's not immediately clear *how*. We do have the diagonal tensor elements (D_{xx}, D_{yy}, D_{zz}) that come with a clear interpretation (*i.e.* the ADC values along x, y and z). From the ADC peanuts in Fig. 2.9, we can tell for instance that the value of D_{xx} will coincide with the maximal ADC value of the peanuts right in the middle of the GCC, because those peanuts are nicely *aligned* along x. However, as we move away from that middle region, the orientation of the peanuts changes, causing D_{xx} to gradually take on lower values. The information on the maximal ADC value of the peanuts is now "spread out" somewhere between D_{xx} and D_{yy}. And to make matters even worse, the information on how this "spread" is balanced between those components, is in turn encoded somehow by D_{xy}, one of those elusive off diagonal tensor elements. That's also why we didn't run into all that trouble right in the middle region: D_{xy} equals zero in that part of the GCC.

2.5.2 Reasoned wishful thinking of alternatives

So, what is at the core of all this confusion and why do we need to be so tedious about trying to infer useful information from the tensor components? The answer is simple: our definition of axes (*i.e.* x, y and z) is in fact quite artificial and – more importantly – *very rigid*. To formulate it in another, maybe more clear, way: the axon bundles simply couldn't care less about how we happened to define our *globally fixed axes*; they just happily twist and curve through the full 3D space. On those rare occasions where the tensor perfectly *aligns* to our predefined axes, we get lucky: the off diagonal elements become zero and the 3 diagonal components describe the *shape and size* of the tensor in a more direct, intuitive manner. But how do we solve our problem in all those other voxels then? As we just stated, the axon bundles are not going to adjust themselves to our axes; and thus the only solution will be to *adjust our axes* to them in each voxel instead. So, what we are looking for is a new description of the diffusion tensor that provides *a set of axes aligned to the tensor* as well

as 3 "*new diagonal tensor elements*" to describe the tensor within this new local set of axes (the "new off diagonal tensor elements" become zero). The benefits are twofold. These "new diagonal tensor elements" should provide us with everything we need to know about the *shape and size* of the tensor, *independently of its orientation*. On top of that, the customized set of axes by itself also describes the full 3D *orientation* of the tensor. To conclude, such a representation thus effectively *splits up* information about the orientation and the shape/size of the tensor, while the classical 6 tensor elements *mix it all up*.

2.5.3 Maths to the rescue: the eigendecomposition

Now that we know what we want, the question remains how to obtain it. In this case, we are lucky: the mathemagician can help us out with something called *eigendecomposition*. Applied to the diffusion tensor, it basically boils down to rewriting the 3×3 symmetric tensor in the following format:

$$\mathbf{D} = \begin{bmatrix} \vdots & \vdots & \vdots \\ \epsilon_1 & \epsilon_2 & \epsilon_3 \\ \vdots & \vdots & \vdots \end{bmatrix} \cdot \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \cdot \begin{bmatrix} \cdots & \epsilon_1 & \cdots \\ \cdots & \epsilon_2 & \cdots \\ \cdots & \epsilon_3 & \cdots \end{bmatrix} \tag{2.10}$$

where $\lambda_1 \geq \lambda_2 \geq \lambda_3$, and ϵ_1, ϵ_2 and ϵ_3 are 3 element unit vectors that are mutually perpendicular to each other. The right hand side intuitively reads: start with a tensor (with diagonal elements λ_1, λ_2 and λ_3) aligned to the axes (x, y and z), and then *reorient* it to a new set of axes (ϵ_1, ϵ_2 and ϵ_3). The process of eigendecomposition aims to *reverse* this set of actions: it starts with the diffusion tensor $\mathbf{D}$, and subsequently tries to figure out which axis aligned tensor could have been reoriented to which new set of axes in order to obtain $\mathbf{D}$. The result is referred to as the *eigenvalues* (λ_1, λ_2 and λ_3) and *eigenvectors* (ϵ_1, ϵ_2 and ϵ_3) of $\mathbf{D}$. They come in so-called *eigenpairs* (*e.g.* λ_2 is paired to ϵ_2): an eigenvalue (*e.g.* λ_2) represents the ADC value of the tensor along the direction of the corresponding eigenvector (*e.g.* ϵ_2). The eigenvector ϵ_1 that is associated with the largest eigenvalue λ_1 is also referred to as the *principal eigenvector*. It plays quite an important role in DTI: due to its orientation along the peak direction of the ADC peanut, it's indicative of *the local direction of the axon bundle*. While the *largest* eigenvalue λ_1 equals the *maximal* value of the ADC peanut, the *smallest* eigenvalue λ_3 represents its *minimal* value.

2.5.4 Understanding the eigenvalues, in practice

Maps of the eigenvalues (λ_1, λ_2 and λ_3) are presented in the top row of Fig. 2.11. A strict ordering ($\lambda_1 \geq \lambda_2 \geq \lambda_3$) is always enforced. As stated before, the combination of all 3 eigenvalues fully encodes the exact total *shape and size* of the tensors (and by consequence, the ADC peanuts) by providing the ADC value along 3 perpendicular axes aligned to the tensors (the eigenvectors). In some regions (*e.g.* the GCC, or the white matter in general) a larger mutual difference between the eigenvalues can be seen as compared to other regions (*e.g.* the CSF). This clearly relates to the differing amounts of anisotropy that we could also see in the ADC peanuts. Because all information on the shape and size of the tensors is stored in the eigenvalues, they will also be the basis for other tensor *measures* that are *independent* of the tensor's *orientation*; but we'll get to that later.

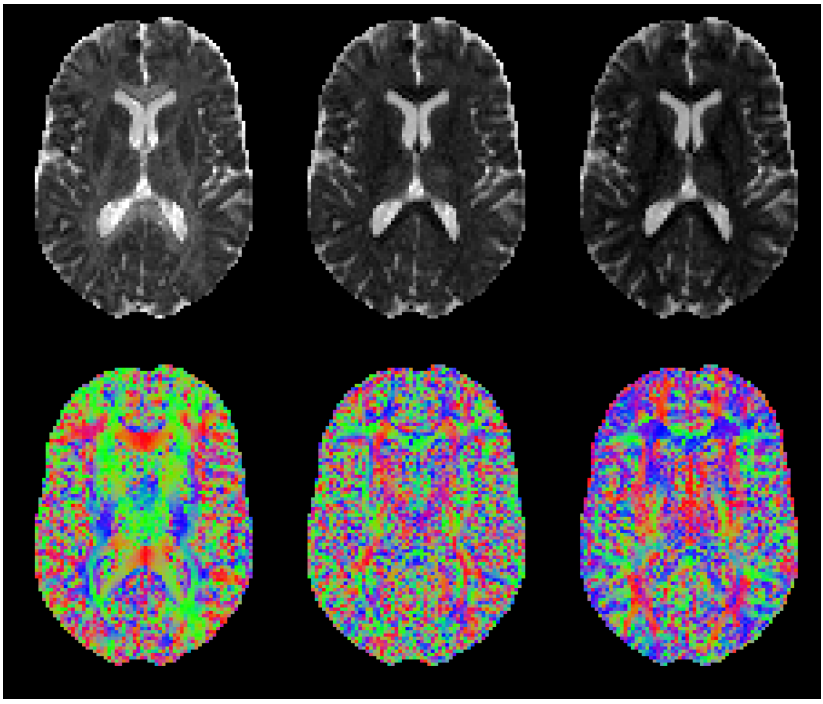

Figure 2.11: *Top row*: maps of the eigenvalues (λ_1, λ_2, λ_3). *Bottom row*: directionally-encoded color (DEC) maps of the eigenvectors (ϵ_1, ϵ_2, ϵ_3).

2.5.5 Understanding the eigenvectors, in practice

Directionally-encoded color (DEC) maps (Pajevic and Pierpaoli, 1999) of the eigenvectors (ϵ_1, ϵ_2 and ϵ_3) are provided in the bottom row of Fig. 2.11. Since each eigenvector has unit length, no magnitude information is represented in these maps; only *orientation* is encoded. This is achieved by assigning the 3 elements of an eigenvector to *the red, green and blue channels* of a color image. As the eigenvector itself is specified relative to the original (x, y and z) axes, the meaning of the colors is similar to the scheme we used before for displaying spherical polar plots: *red* is linked to x, *green* to y and *blue* to z. As mentioned before, one of the most important outcomes of DTI is the orientation of the *principal eigenvector* ϵ_1. Within regions of the white matter (*e.g.* the GCC), the DEC map of ϵ_1 shows a consistent and smoothly evolving pattern that can intuitively be related to the *local orientation of the axon bundles*. In regions such as the CSF, the orientation of ϵ_1 proves to be more or less *random*, resulting in a *noisy* appearance of its DEC map in those particular regions. Associated with the isotropic pattern of diffusion in these regions, we should ideally observe a *spherical* ADC plot (instead of a peanut), satisfying $\lambda_1 = \lambda_2 = \lambda_3$. However, due to random noise in the data, there might be a slight *deviation* from this

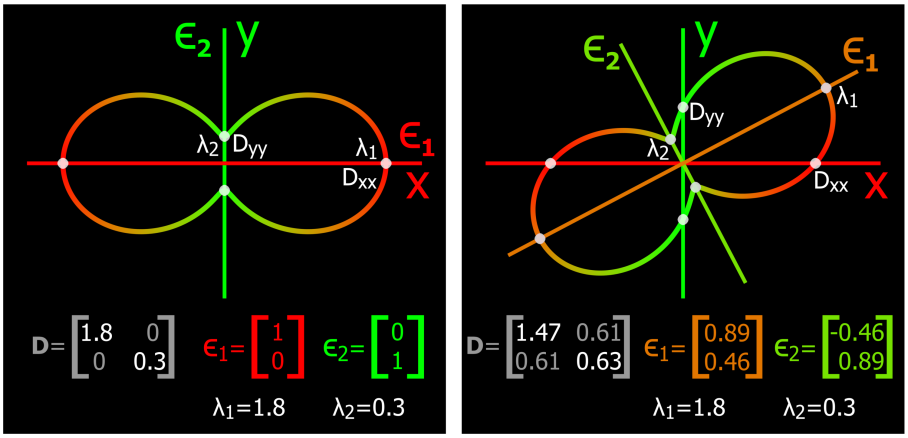

Figure 2.12: Example of 2D tensors and the convenience of eigenvalue decomposition. *Left*: perfectly axis aligned tensor. The diagonal tensor elements directly define the tensor shape. The eigenvectors coincide with the global axes and the eigenvalues are equal to the diagonal tensor elements. *Right*: general (not axis aligned) tensor. The diagonal tensor elements encode the value of the ADC peanut along the global axes, yet do not fully define the tensor shape. The eigenvectors provide a new set of axes along which the eigenvalues directly provide the information on the shape.

pattern. The orientation of the principal (and any other) eigenvector is entirely determined by the *random* noise in such a case. Whenever 2 (or all 3) of the eigenvalues of a given tensor are (nearly) equal, we say that the corresponding eigenvectors become *ill-defined*.

2.5.6 Conclusions

We introduced the *eigendecomposition* of the diffusion tensor in its *eigenvalues* and *eigenvectors*. The 6 diffusion tensor components *mix up* information on the shape, size and orientation of the tensor and it becomes hard to untangle the information we're typically interested in by purely intuitive reasoning on these components. Our new representation, however, nicely *separates* information on the *size/shape* of the tensor from information on its *orientation*. This is achieved by recovering a set of *axes* (the eigenvectors) that is locally aligned to the tensor as well as 3 *ADC values* (the eigenvalues) of the tensor along these new axes. The solution thus comes as a set of *eigenpairs*: a certain eigenvalue encodes the ADC along a specific eigenvector. Whereas the eigenvalues encode the size/shape *independently* of the orientation, the eigenvectors describe the orientation *independently* of the size/shape. A final schematic (2D) example, illustrating these properties and providing an overview of the relation between the most important tensor related numbers we've come across up to this point, is shown in Fig. 2.12. The eigenvector associated to the *largest* eigenvalue is also referred to as the *principal eigenvector*. Mapping eigenvectors is typically done by use of *directionally-encoded color (DEC) maps*. One of the key practices in DTI consists of mapping the *principal eigenvector*, since it is indicative of the *local orientation of the axon bundles*. In regions of white matter such as the GCC, this map shows a *consistent* pattern. In regions of (nearly) isotropic diffusion such as the CSF, however, the principal eigenvector becomes *ill-defined*, leading to a *noisy* appearance of the associated DEC map. Since there are no axon bundles hindering/restricting the diffusion in such a region, the principal eigenvector therein is pretty *meaningless* anyway.

2.6 Visualizations, measures and maps

2.6.1 Aiming for usability: tensor glyphs

While all maps and visualizations (using *e.g.* spherical polar plots) presented up to this point have provided us with great insight in the underlying information that eventually leads to the diffusion tensors and describes their main features,

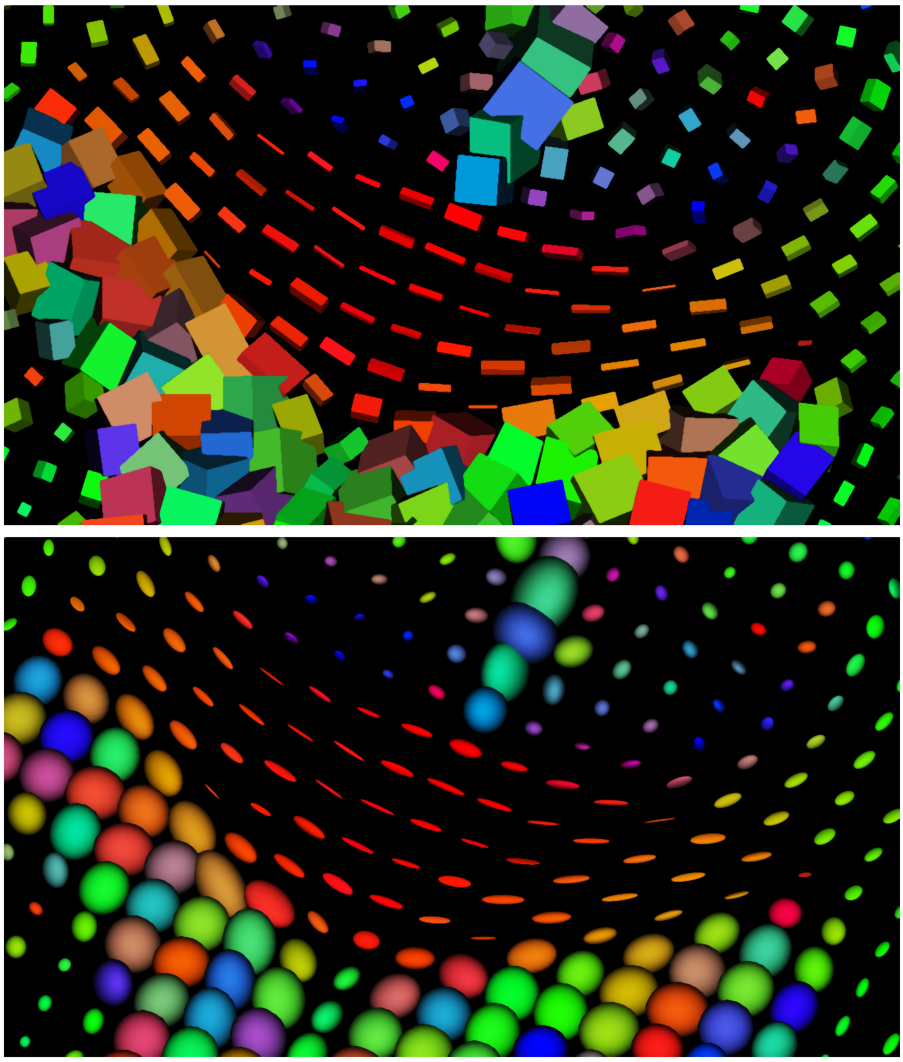

Figure 2.13: Tensor glyphs (*top*: cuboids; *bottom*: ellipsoids) in a region of the GCC. The glyphs are colored according to the DEC map of the principal eigenvector ϵ_1.

we are yet to encounter the visualizations and maps that we're most likely to run into when processing DTI data *in practice*. Let's first have a look at the most common 3D visualization of the diffusion tensor, which is not the ADC peanut we've already become acquainted with (even though it most directly shows all the values that the tensor represents). Instead, we will only visualize the *most prominent features* that define its size/shape and orientation, as provided by its eigenvalues and eigenvectors: meet the *tensor glyphs*! Two variations are shown in Fig. 2.13 for our familiar GCC region: *cuboids* and *ellipsoids*. In general, a mostly primitive 3D shape is chosen (*e.g.* a rectangular cuboid or a scalene ellipsoid) and its 3 main dimensions are *scaled by the eigenvalues* (or a transformation thereof) and *aligned along the eigenvectors*. Optionally, the glyph is *colored* according to the DEC map of the *principal eigenvector* (*i.e.* using the map of ϵ_1 in Fig. 2.11). Comparing our previous ADC peanuts in Fig. 2.9 with the newly obtained glyphs in Fig. 2.13, we can clearly tell the latter score higher on the *usability* scale: they show a more contrasting description of the features that really matter (and still fully define the tensor and thus its ADC peanut). The most commonly visualized glyph shape is the ellipsoid (Basser et al., 1994b), but since a cuboid requires far fewer polygons to be drawn onscreen, it lends itself for faster interaction with larger tensor fields. Interaction with a field of glyphs is useful for a better characterization of their full *3D shapes*: *e.g.* in Fig. 2.13, we freely rotated the slice of glyphs to a certain angle. Especially when the axon bundles are not running "in plane", the ability to freely rotate the tensor field is a helpful addition. A reason to prefer ellipsoids (instead of *e.g.* cuboids) might be that they do not over exaggerate some features of the tensor in cases where those features are not very meaningful or appropriate anyway. A good example is the isotropic diffusion in the CSF, as seen in Fig. 2.13: the cuboids become cubes, but still clearly indicate the orientation of the eigenvectors, even though they are *ill-defined* in this region. The ellipsoids, however, take on the shape of spheres, and thus any visual cues of the eigenvectors inherently *fade away* (apart from the coloration, which is of course also not very informative in this region). Another reason why *ellipsoids* are a meaningful choice, is that they actually come with a true *meaning* when scaled using the *square roots of the eigenvalues* (Basser et al., 1994b): under the model of diffusion that DTI assumes, if we would investigate a single water molecule that starts at the center of the ellipsoid and is allowed to diffuse randomly during a fixed time interval, then there is an equal chance for it to displace to any specific point on the surface of this ellipsoid. It might take a few reads of that sentence before one may grasp its meaning, and we won't even go into why it is true; the fact just is that there exists a pretty good reason to prefer these ellipsoids over any other specific glyph! *In practice*, however, *any glyph will do* for exploring the data (even though some software may offer many different options), as long as it's easy on the eyes and the processing power of

the machine one is working on.

2.6.2 Mapping size: mean diffusivity (MD) and friends

Now let's consider some of the more common diffusion tensor measures. All the measures we're about to present, are so-called *rotationally invariant* measures: they tell us something about the size or shape of the tensors, *independently* of their orientation. Therefore, they are typically defined in function of the *eigenvalues* of the tensors. Let's start with a straightforward one: the *mean diffusivity (MD)* (Basser, 1995). It is defined as follows:

$$\mathrm{MD} = \frac{\lambda_1 + \lambda_2 + \lambda_3}{3} = \frac{D_{xx} + D_{yy} + D_{zz}}{3} \tag{2.11}$$

As simply being the *average of the eigenvalues*, it describes the overall *size* of the tensor and as such represents a rotationally invariant ADC measure. A map of it is provided in Fig. 2.14. The same contrast is sometimes also referred to as the *trace* (Basser, 1995), which equals the *sum* of the eigenvalues. Other related variants exist, such as the pair of *axial diffusivity* (equating to the first eigenvalue) and *radial diffusivity* (equating to the average of the second and third eigenvalues). As seen in Eq. (2.11), the MD can (surprisingly) also be obtained by averaging the *diagonal tensor elements*: even though these individual elements are dependent on the orientation of the tensor, their average is not. An important warning at this point: this does *not* mean that we can simply acquire 3 DWIs using perpendicular gradient directions, and subsequently average the 3 ADCs in order to obtain the *same rotation invariant MD* (Basser, 1995)! It only applies for 3 perpendicular ADC values as evaluated from a *tensor model*, so *6 gradient directions* are still the bare mathematical minimum in order to account for the anisotropy in the measurements!

2.6.3 Mapping fractional anisotropy (FA) and orientation

Next up is the *fractional anisotropy (FA)* (Basser, 1995), probably the most unique selling point of DTI. It is calculated by the following hefty formula:

$$\mathrm{FA} = \sqrt{\frac{3}{2}} \cdot \frac{\sqrt{(\lambda_1 - \bar{\lambda})^2 + (\lambda_2 - \bar{\lambda})^2 + (\lambda_3 - \bar{\lambda})^2}}{\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}} \tag{2.12}$$

where $\bar{\lambda}$ is the average of the 3 eigenvalues. In words, this amounts to the *standard deviation* of the eigenvalues *divided by* their *root mean square*. Or, more simply: a measure for how much the eigenvalues *differ*, but *normalized*

so it becomes *independent* of their absolute magnitude. As such, it describes an aspect of the *shape* of the tensor, *independently* of its size (and of course, orientation). Because of the way the formula is carefully normalized, the FA takes on values in an interval *between zero and one*, the former representing perfect *isotropy* (*i.e.* all eigenvalues are equal) and the latter corresponding to perfect *anisotropy* (*e.g.* the extreme case where only λ_1 would have a nonzero value). An FA map is provided in Fig. 2.14. From this, we learn that the white matter clearly has higher anisotropy than any other (healthy) tissue in the brain. As we already know, in regions of low anisotropy, the principal eigenvector's orientation becomes *ill-defined*. Hence, the FA map is the perfect candidate to weight the DEC map of the principal eigenvector from Fig. 2.11: doing so will *hide* the colors in regions where they are *ill-defined* (*i.e.* where they are noisy, confusing and meaningless). The result is known as the *DEC FA map* (Pajevic and Pierpaoli, 1999), and is also presented in Fig. 2.14. This map isn't the most iconic DTI map for no reason: it's a very handy one and becoming acquainted with the color encoding is key to quickly interpreting a lot of the valuable and unique information in the dataset at once. Mentally processing DEC should become second nature; *red for x (left-right), green for*

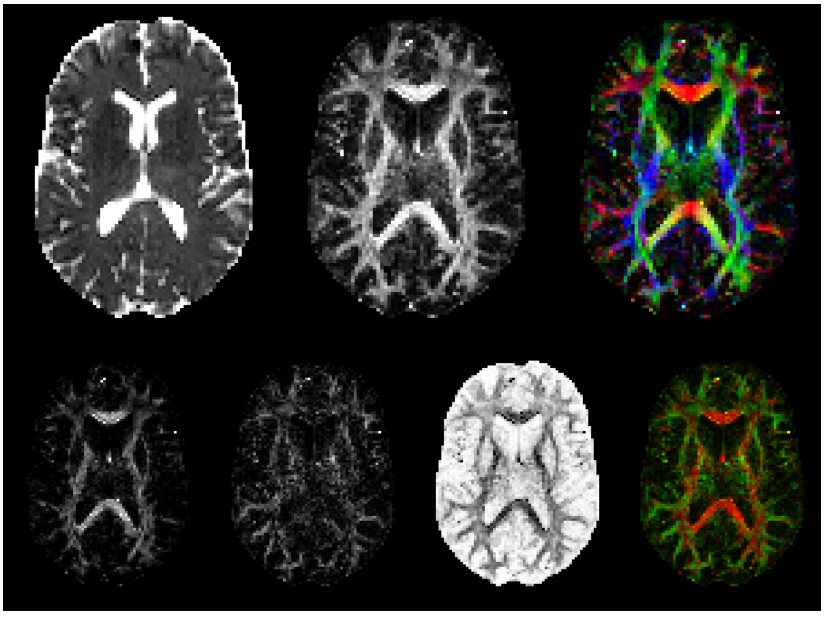

Figure 2.14: *Top row*: maps of the mean diffusivity (MD), fractional anisotropy (FA), DEC FA. *Bottom row*: maps of a linear measure (c_l), planar measure (c_p), spherical measure (c_s), combination of c_l and c_p using red and green color channels.

y (back-front) and blue for z (bottom-top). Take a look again at the original (grayscale) FA map. Notice how easily we could be tempted to believe that every bundle-like feature of this map represents an in-plane axonal bundle. Now shift your attention back to the DEC FA map, and realize that *blue* stands for an orientation *perpendicular* to the visualized slice. There you have it; that's some *indispensable DEC information* for you!

2.6.4 Exploring shape space and reaching beyond...

Finally, let's briefly touch upon a triplet of slightly more exotic measures: a *linear* measure (c_l), a *planar* measure (c_p) and a *spherical* measure (c_s) (Westin et al., 1997). This is what their formulas look like:

$$c_l = \frac{\lambda_1 - \lambda_2}{\lambda_1 + \lambda_2 + \lambda_3} \qquad c_p = \frac{2 \cdot (\lambda_2 - \lambda_3)}{\lambda_1 + \lambda_2 + \lambda_3} \qquad c_s = \frac{3 \cdot \lambda_3}{\lambda_1 + \lambda_2 + \lambda_3} \tag{2.13}$$

All of them are again automatically restricted to an interval of values *between zero and one*. The *sum* of these 3 measures exactly equals *one*. The full triplet of measures provides the coordinates of our tensor in some "*shape space*": a higher c_l means a more linear, *prolate*, cigar-shaped tensor ellipsoid; a higher c_p means a more planar, *oblate*, pancake-shaped tensor ellipsoid; a higher c_s means a more spherical, *isotropic*, ball-shaped tensor ellipsoid. Just like the FA, these measures each describe an aspect of the *shape* of the tensor, *independently* of its size (and of course, orientation). One could even use them to come up with new *anisotropy* measures, such as $1 - c_s = c_l + c_p$ (Westin et al., 1997). Maps of the shape measures are provided in Fig. 2.14. We also present a map where we employ the red and green color channels to encode c_l and c_p. This final map's absolute intensity thus equals the custom anisotropy measure we just mentioned, while the color somehow shows what "*kind* of anisotropy" is present: *linear or planar*. Interestingly, we notice (in the presented slice) that mostly the central part of the corpus callosum (including the GCC region we have been considering in all our examples) shows highly linear behavior, while *many other regions* of white matter contain a decent portion of *planar diffusion*. Reasoning about the axon bundles as a bunch of cylindrical tubes, as we did before, cannot simply cause such a pattern if all axons in the voxel are coherently running along the same direction: they must be curving or dispersing (within certain planes), or maybe more than one population of axons is present in such voxels. Whatever the underlying situation might be in those voxels, the planarity hints at certain *limitations* of the DTI model...

2.6.5 Conclusions

We introduced some of the most *mainstream* visualizations and maps that one is bound to run across in daily DTI practice. These include the visualization of the tensors by *glyphs* (most notably the *diffusion tensor ellipsoid*) and maps of the *mean diffusivity (MD), fractional anisotropy (FA), DEC FA* as well as a slightly more exotic triplet of *linear, planar and spherical measures* (c_l, c_p, c_s). Other measures exist, but these are typically variations of – or at least heavily inspired by – the ones we presented: axial diffusivity, radial diffusivity, the trace, several variants of anisotropy measures, etc... The most *standard* measures of them all, however, are the typical couple of MD and FA: the former representing the average *size* of the tensor (independent of shape and orientation) and the latter encoding its *anisotropy*, an aspect of the *shape* of the tensor (independent of size and orientation). Information on the *orientation* can also be included by *combining* the FA map and a DEC map of the principal eigenvector in order to obtain *the DEC FA map*. This map is not only iconic for DTI, it's also a very handy tool to quickly gain insight in any DTI dataset.

2.7 Tensor Fitting methods

2.7.1 Facing the issue

Up to now, we've been taking an essential step in the whole process for granted: the actual *tensor estimation*. As a matter of fact, this is the *least trivial* step along the pipeline; everything else we've discussed up to this point simply consists of applying some *well-defined* and quite straightforward formulas in the right order. Even the eigendecomposition – or at least what we desire it to yield for an outcome – is exactly defined (*i.e.* Eq. (2.10); no more, no less), and we can rely on computer science to provide us with an algorithm that does the job. In those cases where the outcome proved to be ill-defined (*e.g.* isotropic diffusion), the ill-defined parts of the outcome (*i.e.* the eigenvectors) were not informative anyway. But the *tensor estimation*... that's an entirely different beast! Different, because this time even the definition of "*what we want*" is not all that clear. Or is it? We simply want the tensor to *fit* the data (or the ADC values...?) *as well as possible*, right? But what is "*as well as possible*"? It's *vague*, that's what it is; and hence, a plethora of definitions and associated fitting methods exist. While it easily provides enough material to write a decent book on the subject alone, consider the following as your average quick and dirty hitchhiker's guide to making the right choice. In the foreign restaurant of tensor fitting, it should enable you to more or less translate the menu, pick

something that you're not allergic to, all the while giving you the confidence that your choice will leave you satisfied up to a certain level, but also allow you to leave the restaurant in time so you can still catch your bus. How they actually arrange stuff in the kitchen though, is the least of our concerns. Have a seat; the daily menu consists of: *linear least squares (LLS), weighted linear least squares (WLLS)* and *nonlinear least squares (NLS)*. On top of that, there's also today's special: *robust estimation of tensors by outlier rejection (RESTORE)*. Now let's have a look at our advice!

2.7.2 LLS: quick and dirty

Linear least squares (LLS) is the most basic choice. It will solve your system of equations – of which each single one takes on the form of Eq. (2.9) – by *minimizing the sum of squared residuals* of those equations. In the case where we have more than 6 equations (as opposed to only 6 unknowns), it is typically impossible to perfectly satisfy all equations. The error or difference that still exists between the left and right hand sides of one such equation, is referred to as its *residual*. In practice, minimizing the sum of the squared residuals of all the equations amounts to "*spreading out*" those unavoidable residuals as much as possible *over all of them*. This is equivalent to stating that each equation has an "*equal say*" in the process of the fit. A great advantage of LLS is that it can be implemented as a *single-step process*. That's right: it just takes a single specific operation on the whole system of equations to automatically obtain the solution that optimally minimizes that sum of squared residuals. Depending on your hardware and the size and number of DWIs in the dataset, you can have your tensors rolling out in mere *seconds*! In fact, the tensors we generated in this chapter – and thus all the resulting visualizations and maps we've been looking at – are the result of a single quick application of LLS. For that specific purpose, LLS is certainly well suited: *qualitatively* speaking, our tensors and the subsequently calculated maps of tensor measures look *perfectly fine*. So, why don't we just stick with LLS for all intents and purposes then? The *problem* is subtle and quite well hidden: it actually concerns the fact that we allowed each equation to have *an equal say* in the fit. This assumes that each value that we are trying to fit, was provided to us with an error (*i.e.* the measurement noise) of a "similar magnitude". In more professional words and adapted to how we specifically formulated Eq. (2.9): LLS assumes that the noise *on the ADC values* (*i.e.* the right hand side of each equation, which we are trying to fit with the left hand side) results from a distribution that has the *exact same variance* for all our different ADC values. While this is the case for our DWI values (that originate directly from the scanner, where the noise is "officially" added to the measurements), it does *not* apply to the ADC values: the distribution of

the noise on the original data is logarithmically transformed along with those data to obtain the ADC values! You might remember that we mentioned this before, additionally stating that it would happen to "bug us later on". So now, here it is: officially bugging us. The problem at hand is that each "measured" (*i.e.* calculated) ADC comes with noise of a *different* variance: *i.e.* we can trust some ADCs more (or less) than others. It seems sensible to *weight* the amount of say of each equation in the fit with this information. That's where *weighted linear least squares (WLLS)* kicks in.

2.7.3 WLLS: still quick, less dirty

Weighted linear least squares (WLLS) assigns to each equation (still of the form of Eq. (2.9)) a *weight* according to how much the original noise variation is affected by the logarithmic transform of the data. These weights directly *depend on the magnitudes of the original data* (*i.e.* the intensity of the different DWIs). In practice, once these weights are known, only a limited modification has to be made to LLS to take them into account and obtain a WLLS fit, which now truly provides us with the optimal correct fit! It still only takes a single (slightly bigger) operation on the whole system of equations to get this solution: it might take a few extra seconds, but it still just remains a matter of mere *seconds* to have your tensors again rolling out; yet much more *accurately*. If we would have generated all the maps in this chapter based on a WLLS fit of the tensors, you wouldn't have noticed the difference: it doesn't suddenly change the *visually* informative contrast of those maps. For *quantitative* purposes (such as group studies) though, it certainly matters, *a lot*: WLLS already *removes* a great deal of inherent *biases* on final measures (such as MD and FA) that are typically caused by careless use of LLS. So, why don't we just stick with WLLS for all intents and purposes then? Again, a sneaky problem manifests itself: to determine the weights, we need the magnitudes of the original data... *without the noise*. Of course, once we have obtained a fitted tensor, we could reason that we got rid of the noise (because, optimally, only the noise is left in the residuals). We could then evaluate that tensor for all gradient directions and calculate from the ADCs back to the DWIs, *i.e.* the *noiseless* magnitudes that we needed to determine the weights. So, if we could obtain a fitted tensor, then we would also have our weights; but in order to obtain a fitted tensor, we need those weights in the first place. Yes, that's a *chicken-and-egg problem* we're facing here. No perfect solution exists (and thus, unfortunately, also *WLLS can never be perfect*). A first approach could be to just use the magnitudes of the original noisy data to determine the weights. This may, however, result in a *worse* outcome as compared to using plain old LLS! A second trick is to start by performing *a LLS fit*, and get DWI magnitudes from this fit (where

the noise should then already be accounted for up to a great extent) *in order to determine the weights for a subsequent WLLS fit.* One could then even repeat this process in the hope of getting gradually better fits and subsequent weights for the next fit (but this typically does not add much: most of the "magic" is in using that first LLS just for a robust set of weights). In practice, *it's all still fast*: a LLS followed by a WLLS. The *danger* lies in the fact that some software packages might perform WLLS using the first approach, yielding worse results as compared to LLS. On the other hand, a responsible implementation of WLLS using the second approach should *definitely lead to better results*, without any significant increase in computation time as compared to LLS: it's typically still done in mere *seconds*! However, we can never truly know the correct weights due to the chicken-and-egg format of the problem: so the approach doesn't fix everything. The core of the *original* problem was that the noise got *logarithmically* transformed in the ADC values, and that we formulated Eq. (2.9) based on an ADC value in the left and right hand side of the equation. Knowing now that we actually want to compare the values of the original signals, can't we modify that equation so it compares stuff *that isn't logarithmically transformed*? Easy enough: just remove the logarithm by taking the *exponential* of both sides! Now we are facing the correct form of the equation, but sadly, it also lost its linearity in the unknowns: the linear sum of these unknowns on the left hand side now appears under that *exponential function.* Long story short: it's a nonlinear equation. That's where *nonlinear least squares (NLS)* kicks in.

2.7.4 NLS: a long and brave quest in the mountains

Nonlinear least squares (NLS) will try to solve an overdetermined system of *nonlinear* equations, again aiming to *minimize the sum of squared residuals* of those equations. Going into details about this one is nigh impossible: many methods exist. They all share a common thing, though: they take a much, much *longer time* to reach a solution as compared to LLS and WLLS. They are basically facing the fiendishly difficult problem of finding the lowest point in the lowest valley of a mountainous landscape in a 6 dimensional world. Actually, LLS and WLLS also did, but due to the specific simple shape of the landscape when the equations are linear, they could come up with a nifty trick of finding that lowest point in *a single step.* NLS, on the other hand, is just dropped somewhere on the landscape and has to start a *walk* in the unprepared hitchhikers fashion: without a map (because the landscape is too big and complex) and just relying on its eyes and feeling to *gradually* move to lower regions. In theory, truly solving the NLS problem will yield the optimal result. However, the problem is not easy to solve. Due to the *limited range of sight* in the mountains, a NLS algorithm might get stuck in a *suboptimal* valley

(not knowing there exists another lower valley somewhere). Some algorithms are more *robust* against this than others, but it's nearly impossible to come up with an algorithm that *never* makes these mistakes. In general, many NLS algorithms exist that will in most cases further *outperform WLLS*. In some specifically challenging voxels though, such an algorithm *might fail to converge* or get stuck in the previously mentioned *suboptimal* valleys. If and when the algorithm might detect this, it could for instance perform a WLLS fit instead (still better than nothing or something really wrong, right?). Because NLS algorithms are forced to take a walk in the mountains anyway, they may also come with extra bells and whistles allowing them to generate a solution that specifically satisfies some *constraints*. Due to noise in the data, LLS and even WLLS can sometimes come up with tensors that have one or more *negative eigenvalues*. Of course, this doesn't make sense: negative eigenvalues, and thus negative ADCs, have *no physically sensible meaning*. A NLS algorithm can be guided to not encounter such unwanted cases in the first place: barriers can be put up on the landscape in order to simply deny the NLS hitchhiker access to these forbidden areas. This all typically does come at an extra computational cost, and thus your valuable *time*. Depending on your hardware, the size of the dataset and the kind of NLS algorithm (and the bells and whistles it might come with), some of these strategies may take anywhere from a few *minutes* to several *hours* to finish calculating your tensors. Using a *brain mask* (so no unnecessary calculations are performed for voxels outside of the brain) is typically strongly advised to reduce running time. And guess what? Even if the hitchhiker would be so extremely experienced that he would always find the lowest point in the landscape, his optimal NLS solution could still be unsatisfying. That's because the data aren't only messed up by noise, but possibly also by *outliers*! Motion, distortions, cardiac pulsation, signal dropout, ghosting... *artifacts* are abundant in MRI. Some can be avoided during acquisition, others can be partially dealt with by preprocessing, but in the end some *still leave their mark* on the data when we offer it to our favorite tensor fitting method. They cause *outliers*: data points that have lost all of their informative value by taking on truly silly values that *don't fit the picture*, at all. That's where *robust estimation of tensors by outlier rejection (RESTORE)* kicks in.

2.7.5 RESTORE and beyond: expecting the unexpected

Robust estimation of tensors by outlier rejection (RESTORE) (Chang et al., 2005), as its name suggests, will handle outliers by *rejecting* them. To reject them, they first have to be *detected* though. To do this, it will start with a NLS fit. It will subsequently assign each measurement a *weight*, depending on how well it fits the picture. Another NLS is performed, where each equation is

weighted according to how well its measurement fit the picture before. This process is *repeated until convergence*. The final weights should now be a reliable measure for how well each measurement does (not) fit in, *i.e.* for its "*outlier-ness*". Those measurements that meet a certain threshold are officially regarded as *outliers*, and simply kicked out of the game. The *final fit* is then performed by employing only the surviving "*non-outlier*" data. While this is an ingenious and very *robust* strategy, it does have a few implications. A first one is the fact that it might have to perform *several* subsequent NLS fits: that will surely have an impact on the total computation *time*. It could on average take more than 3 times as long as compared to a single NLS fit (Chang et al., 2005). A second one is the fact that, after kicking out a possibly decent amount of outliers, *enough* measurements should of course still be left to reliably obtain the final fit. Those measurements are even needed to actually reliably classify the other ones as outliers in the first place. Hence, data *redundancy* is an important requirement. Even very recently, further improvements have still been made to relax that redundancy requirement up to a certain extent (Chang et al., 2012). And to top it off, even the most advanced NLS procedures (with or without a mechanism to account for outliers) still aim to *minimize the sum of squared residuals*; which means they inherently assume a *Gaussian distribution* of those residuals. The noise on the MR signal magnitude, however, results in a *Rician distribution*. For *lower SNR* data, such a Rician distribution starts to deviate significantly from a Gaussian one; *i.e.* for DWI data acquired using a *higher b-value*, this adds yet another challenge. Even though the DTI model is about 20 years old now, the fitting problem still remains *far from trivial*.

2.7.6 Conclusions

We took a bite out of some of the most common tensor fitting methods: *linear least squares (LLS), weighted linear least squares (WLLS), nonlinear least squares (NLS)* and *robust estimation of tensors by outlier rejection (RESTORE)*. It is typically said that this specific ordering is one of *increasing complexity*, implying *increasingly better results* at the cost of an even steeper *increase of computation time* (especially for the nonlinear methods). This is generally true; provided that each variant is implemented as well as possible (we rely on the responsibility of the software developers here). If your dataset has enough data *redundancy* (let's say, DWIs for more than 30–40 unique gradient directions (Chang et al., 2012)), we could easily always advise you to use *RESTORE*. However, depending on the specific implementation of RESTORE, the hardware, the size or even number of datasets you have to process, etc... it might take quite a while (possibly up to *several hours*) before you have access to your tensors for further processing. All the bells and whistles in these advanced

nonlinear algorithms may not be necessary, if you're just concerned about having a quick *qualitative* look at the data. For *quantitative* purposes though, we certainly advice to go "*beyond LLS*". A very big gain is already achieved by WLLS (if implemented responsibly), at a minimal extra computational cost. Certainly be on the lookout for the method your favorite piece of DTI software is packing, or even what different choices it might be offering; as you now speak and understand some basic tensor fitting language!

2.8 Final conclusions

In this chapter, we provided an overview that took us all the way from the raw DWI data to the diffusion tensor and even further to some of the more common visualizations and measures. This fact by itself makes for the most important conclusion: while the more "classical" imaging modalities (*e.g.* T1, T2, etc...) are obtained straight from the scanner, the maps that are typically employed in the practice of DTI (*e.g.* MD, FA, DEC FA, etc...) result from a *postprocessing* pipeline: *i.e. these maps are calculated, not directly acquired.* Most of this pipeline is clearly defined; but for the actual *tensor fitting*, there are *quite a few options.* The more advanced methods may also take a reasonable time to be computed. Some *scanner software* offers the option to directly show and export MD, FA, DEC FA and even other maps; however, *don't let that fool you*: this software still has to go through all the steps we've come across in this chapter. Also, if the software almost instantly provides you with *e.g.* a DEC FA map, you should now be aware that it may probably not have performed much more than a simple LLS fit (which might of course be sufficient, if you're just qualitatively inspecting the data). The scanner software also has to rely on the same DWI dataset for this, and thus is *not any more or less reliable* in general than any other piece of software if it comes to providing you with accurate maps: if you do use its features, certainly also try to find out what (tensor fitting) algorithms it employs under the hood! If you want to take advantage of the plethora of different available *(freeware) software packages* that implement several advanced tensor fitting methods (and further postprocessing steps, such as fiber tractography), you'll need to *export* the raw DWI data from your scanner. We've also stressed at several occasions that these images are quite worthless if they don't come with the *accompanying gradient directions and b-values.* More and more manufacturers are starting to take this into account and tuck that information safely away in *e.g.* the DICOM headers, the headers of their own proprietary formats, or even in separate files (containing a *gradient table* in one way or another). However they do it, just try to somehow make sure that it is effectively packed with your data. Your next concern then is to get it *imported* correctly into your DTI software package. Unless that package

supports a whole list of different (more and less) standards, you might be up for yet another daunting task. We're lucky up to some extent, however, as the "*diffusion community*" and the specific supporting communities revolving around some software packages are often very active and responsive: your specific question could be answered quickly after a simple email to a support mailing list. Once you get your workflow up and running, the use of DTI in your daily practice should provide you with *new and exciting insights*!

Chapter 3

Problem statement

3.1 Beyond DTI

Near the end of the previous chapter, we found that a large amount of voxels showed a pattern closer to planar diffusion than linear diffusion. Such a pattern is simply incompatible with a single bundle of axons running perfectly parallel. The problem at hand is in the so-called partial voluming. As mentioned, the current typical voxel size of DWI acquisitions ($2 \times 2 \times 2\,\mathrm{mm}^3$ or even larger) is rather crude, compared to the size of the axons we're trying to obtain information on. It is not unreasonable that, at such a voxel scale, even the complete bundles of axons are involved in more complex patterns. Over the years, the findings from a growing amount of techniques have taught us that we can expect a multitude of configurations within a single voxel, including (but not limited to):

- Crossing: two or more bundles, running at different angles, are (almost) touching each other. Depending on where the voxel grid is exactly positioned relative to the brain, the bundles might contribute to a fraction of the DWI signal in a single voxel. Furthermore, bundles can also be truly interdigitating. In such a case, two or more bundles (running along different directions) pass straight through each other. The axons can be nicely interwoven up to a certain level. If each contributing bundle has the exact same microstructural properties (axon density, diameter, dispersion, myelination, etc...), they would idealy contribute a similar amount to the DWI signal in the voxel. In reality, those properties can easily differ, rendering the situation even more complex.

- Dispersion: the axons are actually anything but exactly straight cylinders. As such, a whole bundle of them doesn't have to consist of exactly parallel structures: either axons are still relatively straight throughout a single voxel but their individual orientations differ (*i.e.* no longer parallel), or they can not even be assumed to be relatively straight. If the distribution features axial symmetry, the DTI model may still indicate a good average direction though.

- Bending: it may seem trivial, but bundles as a whole curve throughout the 3-dimensional space. Due to our crude voxel scale, differences in amounts of curvature will have an impact on the amount of observed dispersion. As diffusion measurements are inherently symmetric, the data in a single voxel doesn't even provide the necessary information to distinguish this setting from a simple dispersion setting. The surrounding voxels might provide hints on the configuration though. Note that bending can be seen as a form of dispersion, but in a more spatially (within the voxel) structured way.

- Fanning: in many places, axons fan out into larger sheet-like structures. At the scale of a single voxel, this will again result in apparent dispersion. The polarity of the fanning cannot be resolved from the information within a single voxel. Again, only the surrounding voxels might possibly provide this information. Just like bending, fanning is a spatially structured case of dispersion.

- Kissing: two bundles may approach each other until they (almost) touch, and then diverge again. This is another example of spatially structured dispersion, mostly resembling a combination of 2 opposite bending configurations.

- Many other configurations can be thought of or even constructed as combinations of any of the above.

A range of various examples of possible configurations is provided in Fig. 3.1. While some of these configurations (bending, fanning, kissing) hint at the limitations of recovering structure from indirect and symmetric diffusion measurements itself, others just show the direct limitations of DTI (crossing, interdigitating) while inspiring the search for better models. As such, many models have been proposed over the years. A continuous discussion is ongoing in the domain on whether the time has come to finally leave the DTI model behind us. As it was recently shown by Jeurissen et al. (2013) that about 90% of the voxels in a typical DWI acquisition contain complex structures that cannot be adequately represented by the DTI model, these issues can no longer simply be ignored; especially not when these techniques are to be used in clinical settings

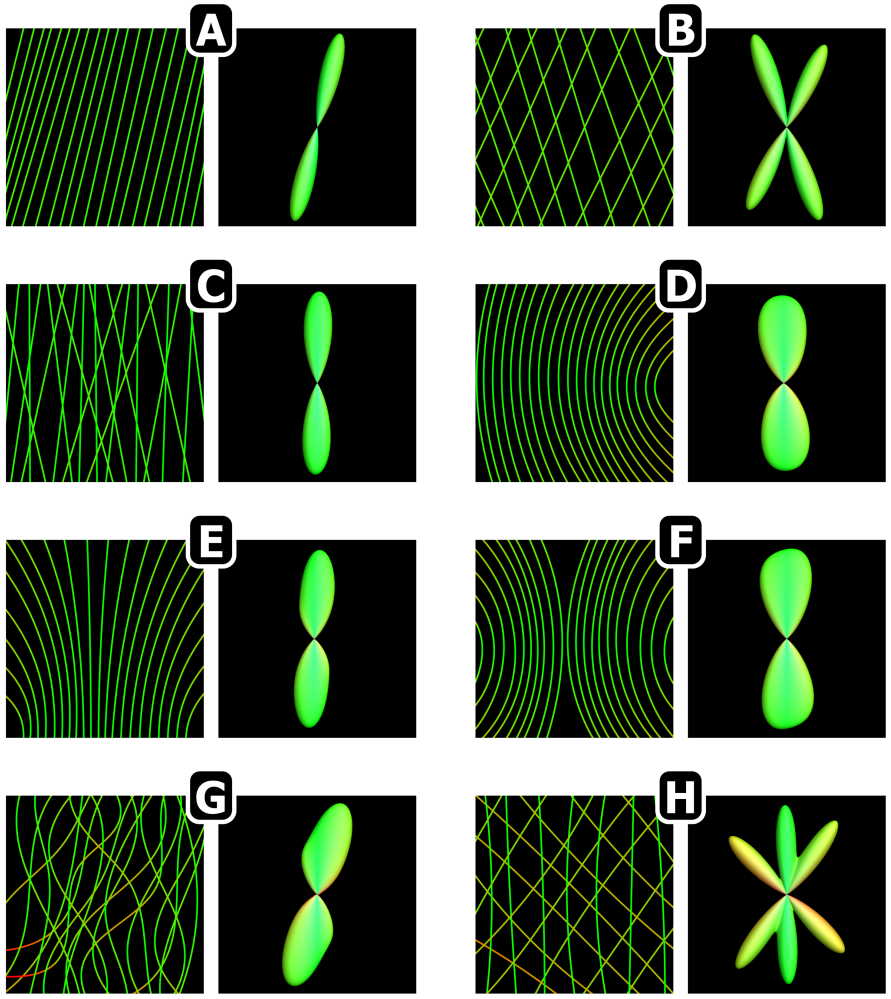

Figure 3.1: Various fiber configurations in 2D; and their orientation distributions. (A) Straight, parallel fibers. (B) Crossing of 2 distinct bundles. (C) Increased random dispersion within a single bundle of straight fibers. (D) A single bending bundle. (E) Fibers fanning out. (F) Kissing of 2 distinct bundles. (G) Dispersion within a single bundle of randomly curving fibers. (H) Crossing of 3 distinct bundles.

where the lives of very real patients are involved. Providing an overview of, or even just enlisting, the bulk of available work on this topic provides enough material for a book in its own right. To anyone interested in specific topics in the general domain of DWI, or other newly starting researchers, I can personally strongly advice the incredible book of Jones (2010). It is an impressively comprehensive work that currently covers every relevant development in this fast-paced domain up to 2010. Specifically to the topic of "crossing fibers" (a general name that is often used for any of the aforementioned complex configurations), a special section was dedicated that contains a complementary couple of chapters by Alexander and Seunarine (2010) and Tournier (2010).

To give the reader a quick but powerful impression of the possibilities that lie beyond DTI though, we simply suggest a look at Fig. 3.2. The top image shows the outcome of a strategy referred to as constrained spherical deconvolution (CSD) (Tournier et al., 2004, 2007). Just as many existing compartment models, it assumes that the signals caused by different fiber populations in a single voxel add up linearly (this basically implies the assumption of no diffusion causing exchange of molecules between the different populations within the measurement time). CSD (and SD in general) is unique, however, in that it does not assume a discrete number of populations (*e.g.* one, two or three; as many compartment models assume). It rather obtains a continuous fiber orientation distribution (FOD) in the angular domain. In such a theory, the diffusion signal can be seen as the convolution of the FOD with the response function that a single fiber would cause. To obtain the FOD, the DWI signal thus has to be deconvolved. To counter the ill-posed nature of such an operation, a non-negativity constraint is enforced on the FOD. These minimal (and fully reasonable) assumptions lead to a result as shown in Fig. 3.2, which already goes miles beyond what DTI can ever achieve. Note especially the appearance of crossing fiber configurations in crucial regions deep within the white matter.

Such a better representation of complex fiber configurations is obviously crucial if one wants to take this a step further. A popular next step is actually shown in the bottom image of Fig. 3.2: fiber tractography. A fiber tractography algorithm, in general, builds up a global continuous structure that matches the local orientational information provided by a model. In this case, the tracks were constructed to match the orientational information of the FOD in each voxel. It is trivial to see that the DTI model would never be able to provide the information to guide the tracks succesfully through the areas containing crossing structures. This is important for one of the most frequently suggested applications of fiber tractography: neurosurgical planning and navigation, where the tracks might provide information on essential structures that should be saved at all costs. Especially for these specific applications, it has been shown that DTI is severely lacking (Farquharson et al., 2013). A succesful example of

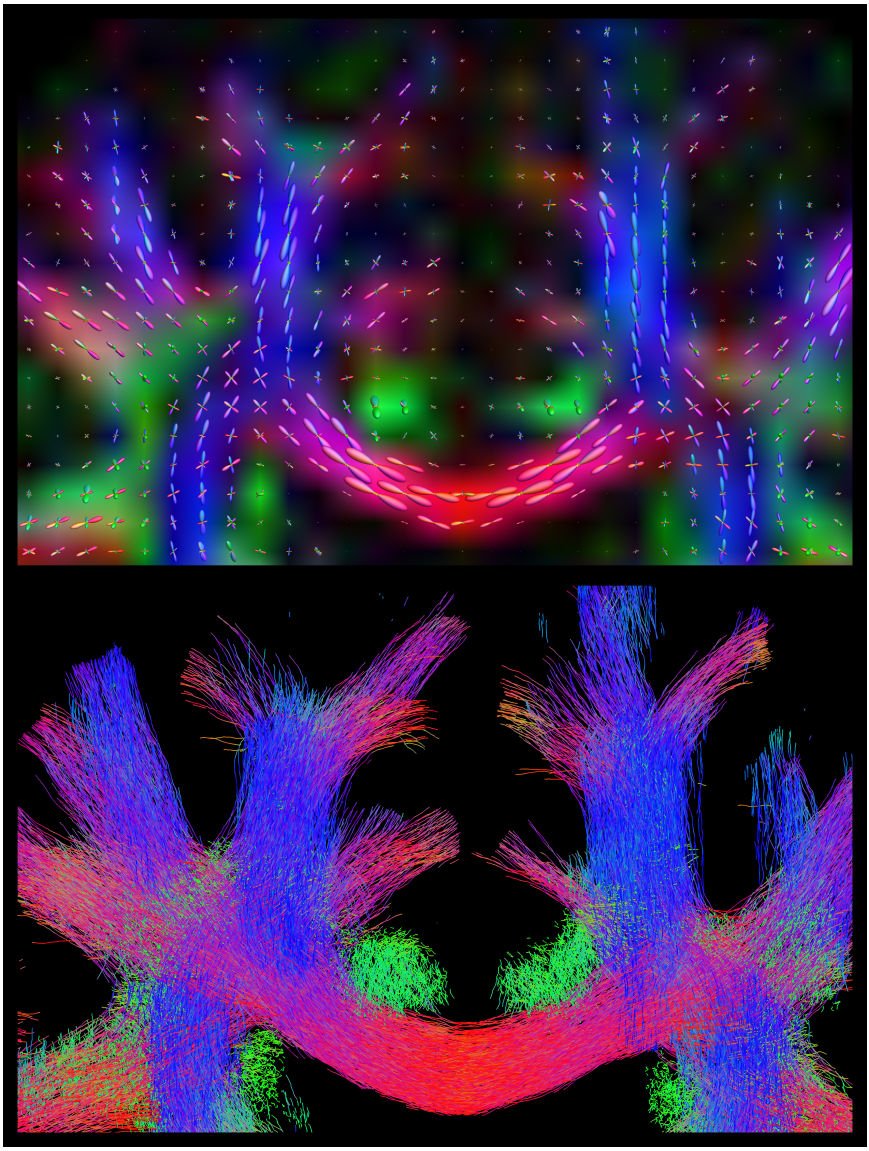

Figure 3.2: *Top:* Fiber orientation distributions (FODs) from constrainted spherical deconvolution (CSD) of the DWI signal, clearly indicating the presence and anatomical plausibility of crossing fibers in crucial regions. *Bottom:* Probabilistic tractography guided by the FODs, maximally exploiting all information on (uncertainty of) directions represented by the FOD.

so called "targetted" tractography, guided by the FODs from CSD, is presented in Fig. 3.3, where I extracted my own corticospinal tract. If I were to be up for brain surgery, I'd definitely want this particular structure to be saved: it directly connects to the rest of my body for most of my sensory input, as well as most of my motoric functions. Employing DTI for this task, would result in large parts of this structure missing.

But even these state-of-the-art techniques should still be used with care and supplemented by other sources of (image) information, as well as careful judging from a professional medical doctor. In the right hands, however, they can become powerful tools that can assist in the complex task of visualization and interpretation of 3-dimensional anatomical image data.

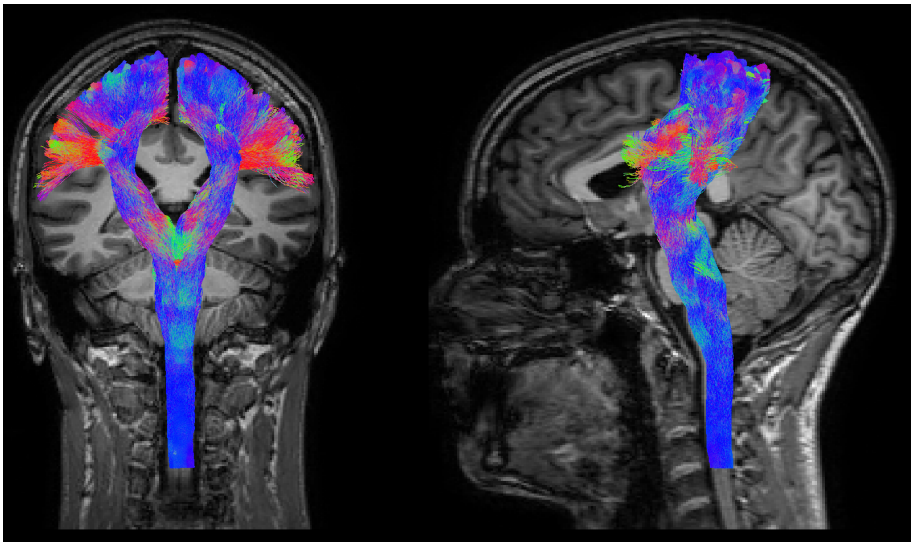

Figure 3.3: So-called "targetted" tractography of the corticospinal tract. By including extra constraints on regions where the tract of interest should run through, entire anatomical structures can be delineated in a largely data-driven manner. This can provide crucial information for *e.g.* surgical interventions.

3.2 This thesis

In this PhD thesis, we intended to design certain representations of (information extracted from) DWI data that take into account the aforementioned variety of complex geometrical configurations. As there already existed a multitude of models and representations in the general domain of DWI at the start of this PhD research, the focus of this work was to track down their limitations for solving relevant other problems that are typically inherent to the practice of medical image analysis (*e.g.* segmentation and registration). Our proposed novel representations aim to offer a greater flexibility that should inherently render many of these problems trivial; yet make as little assumptions as possible on the nature of the data or the properties of the underlying structures. As an interesting by-product, such representations might provide us with greater insight in otherwise complex properties of DWI data and their derivatives. As such, they might even inspire us to develop novel theories on the meaning and interpretation of existing concepts, and allow us to use them for entirely new goals altogether.

An initial challenge was found in the so-called reorientation problem: a challenge unique to DWI data when attempting to apply a spatial transformation. No correct strategies were available yet to solve the problem correctly for the raw DWI signals in q-space, and we could even prove an existing attempt to be fundamentally wrong. While simple rigid reorientation strategies exist for the diffusion tensor model, a requirement to tackle this problem adequately and with a minimum of assumptions on the geometrical configurations found in each voxel, lies in also solving the aforementioned partial voluming problem up to a great extent. A suitable representation of DWI signals for this purpose should separate the contributions of several tissue types that are to be treated differently, as well as separate the geometrical structure of white matter fibers from their actual fractions and signal responses that represent other microstructural properties. Once geometry is separated, its manipulation should become trivial. An extra challenge lies in adequately treating multi-shell DWI data (acquired with different diffusion weightings). As a minimum requirement, the developed method should at least be able to handle multi-shell data (correctly); but the wealth of information in such data might also be used to our advantage. Our final findings went way beyond the original reorientation challenge itself, as uncovering the partial volume fractions results in many other advantages and provides far-going insights on transformational invariant properties of DWI data. The outcome of this line of research includes a novel multi-shell multi-tissue (MSMT) representation and an accompanying novel preservation of principal volume fractions (PPVF) retransformation strategy; and is fully described in Chapter 4.

The second challenge of this PhD was to design a more tangible representation for complex fiber track distributions typically resulting from state-of-the-art probabilistic tractography algorithms. For tractograms, there lies no challenge in applying spatial transformations (this operation is trivial); but rather in managing their complexity for other problems such as registration or segmentation. During the course of the PhD, the technique of track-density imaging (TDI) (Calamante et al., 2010) was introduced. TDI suggests a simple mechanism (counting tracks in a grid of voxels) to map the spatial distribution of tracks. TDI was originally promoted for its so-called super-resolution properties. As a first encounter with the technique, we investigated the variability in appearance of a classical TDI map as resulting from varying noise realisations at the level of the DWI data. The outcome of this simple but powerful experiment is described in Chapter 5. Next, we designed an angular extension of TDI, called track orientation density imaging (TODI). While the resulting track orientation distribution (TOD) representation should allow for better solutions to the aforementioned problems, it also provided us with new insights on complex track distributions. More specifically, it allowed us to understand and explain the amplitude of a short-tracks TOD (where the maximum track length is severely limited) by a novel theory of track-like local support (TLS). Intriguingly, while the technique gets its crucial information from (short-tracks) tractography, it might also provide interesting novel information specifically to guide tractography in its own right. We explored the possibility of such TOD-based tractography and concluded that the short-tracks TOD contains the necessary information to guide tracks along directions that are more likely to correspond to continuous structure over a longer distance; *i.e.* track-like structure! Again, this line of research went far beyond the initial research questions and expectations. Our findings are described in Chapter 6.

Finally, a summary of the main contributions resulting from these works is provided in Chapter 7, along with a multitude of opportunities for future research.

Chapter 4

MSMT representation and PPVF retransformation

Preliminary results on reorientation strategies were presented on two occasions (respectively as an oral presentation and an educational poster):

- Spatial Transformations of High Angular Resolution Diffusion Imaging Data in Q-space. (CDMRI workshop at 13th international conference on MICCAI; Beijing (China), 2010)
- Methods for Reorienting and Retransforming Diffusion Weighted Imaging Data. (19th annual meeting of the ISMRM; Montreal (Canada), 2011)

Further work on obtaining a multi-shell DWI template was presented (orally) on two occasions:

- Constructing a Hybrid Diffusion Imaging Atlas in Q-space. (3th annual meeting of the ISMRM Benelux Chapter; Hoeven (Netherlands), 2011)
- A Hybrid Diffusion Imaging Atlas in Q-space. (19th annual meeting of the ISMRM; Montreal (Canada), 2011)

For the former presentation, an award was also received:

- **Best Presentation Award (1st place)** for "Constructing a Hybrid Diffusion Imaging Atlas in Q-space." (3th annual meeting of the ISMRM Benelux Chapter; Hoeven (Netherlands), 2011)

This work was furthermore extended and presented as a poster:

- Feasibility and Advantages of Diffusion Weighted Imaging Atlas Construction in Q-space. (14th international conference on MICCAI; Toronto (Canada), 2011)

Due to the popularity of the earlier oral presentations and the presence of the work as a poster on MICCAI, I was also invited on the spot at the CDMRI workshop to provide the oral presentation once more:

- A Hybrid Diffusion Imaging Atlas in Q-space. (replacing/invited talk for CDMRI workshop at 14th international conference on MICCAI; Toronto (Canada), 2011)

The following chapter describes the latest advances and rounding up of this line of work, combining several new insights I gained throughout the course of the PhD. This final work is being prepared to submit for review and publishing in a journal.

4.1 Introduction

The domain of diffusion weighted imaging (DWI) has come a long way since its initial development in the mid-1980s (Le Bihan and Breton, 1985; Le Bihan et al., 1986). Over the years, we have gained a better understanding of the accompanying techniques and necessary processing steps involved, furthermore leading to a wealth of new insights in the complex workings of the (human) brain (Jones, 2010). The introduction of diffusion tensor imaging (DTI) (Basser et al., 1994a,b; Basser, 1995) has played a crucial role in this process, as it provided the first model intended to deal with anisotropic diffusion; a particular feature observed in the white matter (WM), as opposed to the other most common "tissue" types found in the human brain, *i.e.* the gray matter (GM) and cerebrospinal fluid (CSF). Measures, such as fractional anisotropy (FA) (Basser, 1995), were proposed to even quantify the presence of this property.

Dealing with anisotropy brings its own challenges, though. One such complication, as compared to "classical" scalar images, arises when applying spatial transformations to these data. Because a spatial transformation may change the orientation of certain structures which are the underlying cause of the anisotropic diffusion, the data or models representing this diffusion have to be adapted in order to keep the angular information consistent with its spatial counterpart. Specifically for DTI, it was suggested to apply a rotation to the diffusion tensor (Alexander et al., 2001). The most advanced strategy, referred to as preservation of principal direction (PPD), takes into account the local (spatial) transformation as well as the original orientation of the tensor to determine the appropriate rotation.

Nowadays, however, the limitations of DTI are well known. The diffusion tensor lacks the ability to accurately represent multiple, or otherwise complex, WM fiber populations in a single voxel. This is a major issue, as it has been shown that about 90% of the WM voxels in a typical DWI acquisition of the human brain contain such complex configurations (Jeurissen et al., 2013). Strategies to overcome this severe limitation typically require a high angular resolution diffusion imaging (HARDI) (Tuch et al., 2002) acquisition.

This of course also results in a next level of challenges related to correctly transforming complex diffusion models. As *e.g.* a shearing transformation can change the angle between multiple WM fiber populations within a single voxel, a simple reorientation strategy will no longer be sufficient. Hence, a complex *retransformation* strategy is required. Such a strategy specifically aimed at 4th order tensors was proposed by Barmpoutis et al. (2007). A more general approach for orientation density functions (ODFs) was proposed by Hong et al. (2009). This method transforms the sampling vectors of an ODF according to

the complete local affine transformation, and also adjusts the ODF samples to preserve the (local) integral of the ODF. A closed-form solution for the latter definition was provided by Du et al. (2012). Another method for the retransformation of fiber orientation distributions (FODs) was proposed by Raffelt et al. (2009, 2012a). This method aims for the same properties as Hong et al. (2009), but achieves it by modeling the FOD as a sum of apodized spherical harmonics (SH) delta functions along a dense set of directions. For the retransformation step, the complete local affine transform acts on the dense set of directions, while the weights are preserved. This method was also successfully combined with a symmetric registration strategy (Raffelt et al., 2011) and incorporated in a framework for quantitative comparison of FODs between (groups of) subjects (Raffelt et al., 2012b).

As a plethora of models exists in the general domain of DWI, a new retransformation definition and accompanying unique strategy would have to be developed for each single one of them. This doesn't seem very practical, and it would furthermore be difficult to make such strategies completely consistent with each other; *e.g.* it can already be trivially understood that the PPD reorientation method for tensors is not *consistent* with any of the previous proposals for higher order models (although the method of Raffelt et al. (2012a) can be seen as a natural *extension* of PPD to FODs). A solution to this problem would be to transform the DWI measurements in q-space themselves. Any model could then (after transformation) still be fit to the retransformed q-space signals. The resulting fits all rely on the same retransformed data, and are as such consistent among each other. A first proposal for the retransformation of DWI signals in q-space was provided by Tao and Miller (2006), and the exact same strategy was also applied in Yap et al. (2010). Both simply propose to directly apply the local forward affine transformation to the gradient directions, and leave the q-space samples themselves untouched. However, we have shown ourselves in Dhollander et al. (2010) that this strategy is *fundamentally wrong*. The key to understanding this lies in the fact that the q-space signals typically show large magnitudes along gradient directions *perpendicular* to the WM fibers. The affine transformation is to be applied to the fibers, however, and not to the gradient directions; as the latter will break the perpendicular relationship. A simple fix that does keep this relation intact is to use the inverse transpose of the local forward affine transformation (Dhollander et al., 2010). While this qualitatively seems to solve the issue, it is not consistent with the aforementioned preservation of (local) integrals that should be applied for ODFs (Hong et al., 2009; Raffelt et al., 2012a). Therefore, we proposed a new method in Dhollander et al. (2010) that correctly preserves anisotropic volume fractions (AVFs). This is basically the direct equivalent of the method of Raffelt et al. (2012a), translated to q-space; obtained by using a single fiber response function instead of a SH delta function. We also showed, however, that this strategy still has a certain flaw:

when *e.g.* the isotropic q-space signal of CSF is sheared, preservation of AVFs results in the appearance of anisotropy (this observation was confirmed in Du et al. (2012) for ODFs). In line with the interpretation of DWI data with regard to tissue structure, this makes it seem as if anisotropic tissue appeared "out of nowhere"; which is obviously an undesirable effect. As a final solution, we proposed to add an isotropic volume fraction (IVF) to our signal representation (Dhollander et al., 2010), that should represent the contributions of isotropic tissues. The final method preserves AVFs *and* IVFs. A brief overview of the aforementioned evolution of methods is provided in Dhollander et al. (2011b). We also applied our retransformation strategy to obtain the first average DWI signal template for multiple shells in q-space (Dhollander et al., 2011a,c).

A further adaption of our original method was proposed by Yap and Shen (2013). Even though our method did foresee an IVF in the model, the naive least squares fit did not feature an intelligent mechanism guaranteeing that only CSF (and GM) would be captured in the IVF. The method proposed in Yap and Shen (2013) added a non-negativity constraint and an optimization towards a sparse fit to the procedure. While this results in a significant extra computational demand (a non-linear fitting procedure in each voxel), it is expected to better distinguish anisotropic from isotropic tissues for single-shell DWI data. It does, however, directly rely on the assumption that the representation of the WM in the model is of an actual sparse nature. This is clearly not the case for fanning or bending structures, or indeed any other form of dispersion beyond what is directly captured in the (fixed) single fiber WM response function. Orientation dispersion (OD) maps from the recently proposed neurite orientation dispersion and density imaging (NODDI) model (Zhang et al., 2012), clearly show that varying amounts of OD are to be expected, even within the WM. The NODDI model disentangles the mixed contributions of neurite density (ND) and OD to the FA. It is shown in Zhang et al. (2012) that this disentangled information can only be obtained in a truly data-driven manner from multi-shell data.

In this work, our aim is to provide a generalized representation of multi-shell q-space DWI data that renders retransformation trivial and allows for increased insight in the properties of such an operation. Therefore, we propose a multi-shell multi-tissue (MSMT) approach to represent the q-space DWI signal by a set of discrete weights; a single weight for each IVF in the model, and a large set of weights corresponding to a dense set of directions for each AVF. The specific MSMT variant employed for this work consists of an IVF for CSF, an IVF for GM, and an AVF for WM. The response functions for each tissue class are obtained directly from the data itself (without any microstructure model assumptions). In line with the findings from Zhang et al. (2012), we will also use multi-shell information (for the data and the repsonse functions) to distinguish the necessary fractions in a data-driven manner. Since our representation

separates the contribution of each tissue class to the signal, each fraction can be treated appropriately during the subsequent retransformation: the WM populations can be rearranged, the CSF stays isotropic (no sudden appearance of anisotropic tissue), and so does the GM (we don't want to deform the individual neuronal cell bodies). This generalized way of retransforming DWI data will be referred to as preservation of principal volume fractions (PPVF), as it is a natural generalization of the PPD reorientation strategy to each separate volume fraction (where the WM consists of a large set of fractions). The MSMT representation's parameters themselves are also much more meaningful and interpretable as compared to some of the classic DTI parameters, such as FA. Furthermore, the discrete directional information represented by MSMT lends itself to a better definition of a directionally-encoded color (DEC) map. We will therefore also propose a novel DEC WM map, that is specifically designed to tackle the most important flaws of the classic DEC FA map. As a proof-of-concept, we will go through the entire pipeline of fitting our MSMT representation (including obtaining the response functions), interpretation of the fitted outcome, retransformation using the PPVF strategy for deformations obtained from a groupwise registration, and finally the construction of an average multi-shell q-space template. The final discussion is dedicated to a brief brainstorm on some properties of our MSMT representation combined with the PPVF mechanism, as well as the opportunities for future research.

4.2 Materials and Methods

4.2.1 Multi-Shell Multi-Tissue (MSMT) representation

We propose a specific MSMT representation of the DWI signal that includes two IVFs (CSF and GM) and one AVF (WM). For each voxel, the DWI signal S in function of gradient direction u and b-value, is represented as a sum of weighted response functions by

$$S(u,b) = f_{\mathrm{CSF}} \cdot R_{\mathrm{CSF}}(b) + f_{\mathrm{GM}} \cdot R_{\mathrm{GM}}(b) + \sum_{d=1}^{D} f_{\mathrm{WM}}^{d} \cdot R_{\mathrm{WM}}^{z_d}(u,b) \qquad (4.1)$$

where $R_{\mathrm{CSF}}(b)$ and $R_{\mathrm{GM}}(b)$ are isotropic CSF and GM response functions respectively, and each $R_{\mathrm{WM}}^{z_d}(u,b)$ denotes an axially symmetric anisotropic single fiber WM response function that has been reoriented to have its central axis of symmetry line up along direction z_d. In this work, we employ a fixed large set ($D = 300$) of uniformly distributed directions z_d, obtained by electrostatic repulsion (Jones et al., 1999a), as shown in Fig. 4.1. Directions, such as u and z_d, can be expressed by a pair (θ, ϕ) of spherical coordinates (or a unit vector of

three Cartesian coordinates). We will refer to the WM response function with its central axis of symmetry along the $\theta = 0$ direction as $R_{\mathrm{WM}}(\theta, b)$. While the total CSF and GM fractions are denoted by f_{CSF} and f_{GM} respectively, the WM fraction is spread out over several directions. The total WM fraction is simply defined as:

$$f_{\mathrm{WM}} = \sum_{d=1}^{D} f_{\mathrm{WM}}^{d} \tag{4.2}$$

In what follows, we define the total of all fractions as

$$f_{\mathrm{total}} = f_{\mathrm{CSF}} + f_{\mathrm{GM}} + f_{\mathrm{WM}} \tag{4.3}$$

and normalized fractions relative to this total are furthermore denoted by

$$\breve{f}_{\mathrm{CSF}} = f_{\mathrm{CSF}} / f_{\mathrm{total}} \tag{4.4}$$

$$\breve{f}_{\mathrm{GM}} = f_{\mathrm{GM}} / f_{\mathrm{total}} \tag{4.5}$$

$$\breve{f}_{\mathrm{WM}}^{d} = f_{\mathrm{WM}}^{d} / f_{\mathrm{total}} \tag{4.6}$$

$$\breve{f}_{\mathrm{WM}} = f_{\mathrm{WM}} / f_{\mathrm{total}} \tag{4.7}$$

of which the triplet $(\breve{f}_{\mathrm{CSF}}, \breve{f}_{\mathrm{GM}}, \breve{f}_{\mathrm{WM}})$ consequently sums to 1.

Provided $R_{\mathrm{CSF}}(b)$, $R_{\mathrm{GM}}(b)$ and $R_{\mathrm{WM}}(\theta, b)$ are known (*e.g.* given by a model, or measured from the data), the fractions f_{CSF}, f_{GM} and f_{WM}^{d} in Eq. (4.1) can be recovered from a set of measurements of $S(u, b)$ by solving a linear system of the form

$$\mathbf{R}\mathbf{f} = \mathbf{s} \tag{4.8}$$

that can be nicely structured as

$$\begin{bmatrix} \mathbf{R}_{\mathrm{CSF},1} & \mathbf{R}_{\mathrm{GM},1} & \mathbf{R}_{\mathrm{WM},1} \\ \vdots & \vdots & \vdots \\ \mathbf{R}_{\mathrm{CSF},B} & \mathbf{R}_{\mathrm{GM},B} & \mathbf{R}_{\mathrm{WM},B} \end{bmatrix} \begin{bmatrix} \mathbf{f}_{\mathrm{CSF}} \\ \mathbf{f}_{\mathrm{GM}} \\ \mathbf{f}_{\mathrm{WM}} \end{bmatrix} = \begin{bmatrix} \mathbf{s}_1 \\ \vdots \\ \mathbf{s}_B \end{bmatrix} \tag{4.9}$$

in case of typical multi-shell data. In $\mathbf{s}$, each $\mathbf{s}_i$ is a column vector containing measurements of the i-th shell $S(u, b_i)$ along several directions. Each $\mathbf{R}_{\mathrm{CSF},i}$ (or $\mathbf{R}_{\mathrm{GM},i}$) has a matching size and is filled with the single constant $R_{\mathrm{CSF}}(b_i)$ response amplitude (or the $R_{\mathrm{GM}}(b_i)$ amplitude respectively). Each $\mathbf{R}_{\mathrm{WM},i}$, on the other hand, consists of D columns that each contain the amplitudes of a single $R_{\mathrm{WM}}^{z_d}(u, b_i)$, evaluated along the directions corresponding to the $S(u, b_i)$ measurements. For the fraction column vector $\mathbf{f}$, we simply have $\mathbf{f}_{\mathrm{CSF}} = f_{\mathrm{CSF}}$ and $\mathbf{f}_{\mathrm{GM}} = f_{\mathrm{GM}}$, while $\mathbf{f}_{\mathrm{WM}}$ contains the D fractions f_{WM}^{d}.

4.2.2 MSMT fitting and regularization priors

While our choice to use a fixed large set of D uniformly distributed directions leaves a simple linear system of equations to be solved, the system in Eq. (4.9) will typically be underdetermined for current realistic acquisition schemes. However, because f_{CSF}, f_{GM} and each f^d_{WM} represent amounts of real tissue classes, none of them should – realistically speaking – take on negative values. Moreover, we expect the *presence of tissue populations* to be sparse in each voxel. At this point, it is critical to state that we would still like to allow (combinations of multiple) complex individual populations such as bending or fanning WM at the scale of a single voxel; *i.e.* varying amounts of *dispersion* of single populations exist, and should not be explicitly discouraged *per se*. In other words, while we could encourage sparse presence of tissue populations, we do not want to encourage sparse *appearance* of any single one of them *per se*. Finally, we'd also like to obtain a somewhat coherent pattern of f^d_{WM} fractions in the angular domain, not an erratic one. This should manifest itself as a large *constant* zero f^d_{WM} fraction region which is not interrupted by single relatively low spurious fractions; and zero, one, or more *coherent* WM populations, each with a possible angular region of dispersion and not interrupted by spurious zero fractions. In this particular work, we will try to include prior information on the aforementioned desirable qualities of the tissue fractions by adding

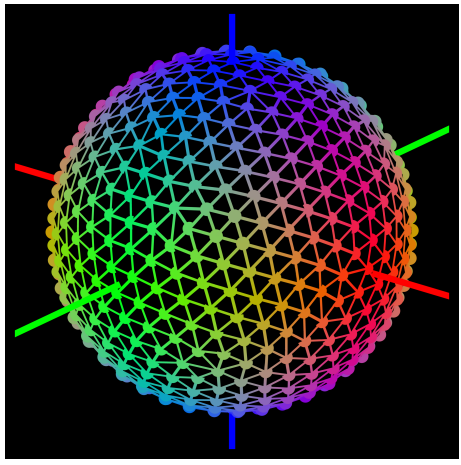

Figure 4.1: The set of 300 directions employed in this work, visualized by 600 points on the unit sphere (*i.e.* 300 symmetrically opposing pairs). The Delaunay triangulation connects "neighboring" directions, for the purpose of defining the corresponding Laplacian matrix. Each direction also displays its directionally-encoded color (DEC). The three Cartesian axes are shown in red, green and blue.

two extra Tikhonov regularization terms to a simple linear least-squares fit of $\mathbf{Rf}$ to $\mathbf{s}$. Furthermore, our choice of regularization terms should resolve the underdetermined nature of the original problem altogether.

The resulting minimization problem is formulated as

$$\hat{\mathbf{f}} = \min_{\mathbf{f}} \left\{ \|\mathbf{Rf} - \mathbf{s}\|_2^2 + \lambda^2 \|\mathbf{Lf}\|_2^2 + \kappa^2 \|\mathbf{Kf}\|_2^2 \right\} \tag{4.10}$$

where the first term is the data-driven linear least-squares fit, the second term minimizes the magnitude of those fractions that are assumed to be zero, and the third term minimizes erratic changes in the angular pattern of the WM fractions. The parameters λ and κ control the (relative) weighting of the regularization terms. The global optimum $\hat{\mathbf{f}}$ can be found by the closed-form expression

$$\hat{\mathbf{f}} = \left(\mathbf{R}^\mathsf{T}\mathbf{R} + \lambda^2\mathbf{L}^\mathsf{T}\mathbf{L} + \kappa^2\mathbf{K}^\mathsf{T}\mathbf{K}\right)^{-1} \left(\mathbf{R}^\mathsf{T}\mathbf{s}\right) \tag{4.11}$$

in a single step.

However, since we do not know *a priori* which fractions could be assumed to be zero, we take on an iterative approach to determine the exact contents of regularization matrix $\mathbf{L}$ and the accompanying final $\hat{\mathbf{f}}$. First, an initial solution $\hat{\mathbf{f}}$ is obtained by setting $\mathbf{L}$ to the zero matrix, *i.e.* effectively dropping this regularization term. For every subsequent iteration, $\mathbf{L}$ is determined in function of the outcome $\hat{\mathbf{f}}$ of the previous iteration. This process is repeated until convergence. The overall layout of matrix $\mathbf{L}$ itself is structured as

$$\mathbf{L} = \begin{bmatrix} \mathbf{L}_{\text{CSF}} & 0 & 0 \\ 0 & \mathbf{L}_{\text{GM}} & 0 \\ 0 & 0 & \mathbf{L}_{\text{WM}} \end{bmatrix} \tag{4.12}$$

where $\mathbf{L}_{\text{CSF}}$ and $\mathbf{L}_{\text{GM}}$ consist of a single value, while $\mathbf{L}_{\text{WM}}$ is a $D \times D$ diagonal matrix. The general idea is to set the diagonal elements $\mathbf{L}_{i,i}$ to a certain non-zero value for corresponding fractions (resulting from the previous fit) that are either negative or very small relative to the total. We have to be careful though, as f_{CSF} and f_{GM} are of a different order of magnitude compared to any single f^d_{WM}, due to the fact that f_{WM} is spread out over D directions. The key to figuring out the appropriate balance for the non-zero values $\mathbf{L}_{i,i}$ as well as the different thresholds to decide upon such a non-zero value, lies in considering the situation for a model where either the CSF and GM fractions are also both spread out over D directions or a model where the WM fraction is also respresented by a single number (*i.e.* f_{WM}); as in such scenarios, the aforementioned orders of magnitude of the fractions don't differ. We start by defining the diagonal elements of $\mathbf{L}_{\text{WM}}$ as

$$\mathbf{L}_{\text{WM},i,i} = \begin{cases} 1 & \text{if } f^i_{\text{WM}} < \tau \cdot f_{\text{total}} / (3D) \\ 0 & \text{otherwise} \end{cases} \tag{4.13}$$

where τ represents a certain fixed fraction (in this work, we always set $\tau = 0.1$). To reason on the threshold for a single f^i_{WM}, one can think of the scenario where the CSF and GM fractions are also spread out over D directions. In such a scenario, there are $3 \cdot D$ fractions and $f_{\mathrm{total}}/(3D)$ thus represents the average of all fractions. Consequently, Eq. (4.13) boils down to: "if f^i_{WM} is smaller than a fraction τ of the average directional tissue fraction, it is considered a valid candidate to be assumed zero in the fit; which is implemented by setting $\mathbf{L}_{\mathrm{WM},i,i}$ to 1". Note that the threshold condition is equivalent to demanding that $\breve{f}^i_{\mathrm{WM}} < \tau/(3D)$. Next, we define the value of $\mathbf{L}_{\mathrm{CSF}}$ as

$$\mathbf{L}_{\mathrm{CSF}} = \begin{cases} 1/\sqrt{D} & \text{if } f_{\mathrm{CSF}} < \tau \cdot f_{\mathrm{total}}/3 \\ 0 & \text{otherwise} \end{cases} \tag{4.14}$$

in order to maintain an appropriate balance relative to the definition of $\mathbf{L}_{\mathrm{WM}}$. The threshold to assume f_{CSF} being zero, is again established at a fraction τ of the average tissue fraction; only this time one has to imagine a scenario where each of the 3 tissue fractions is represented by a single number: the average fraction thus equals $f_{\mathrm{total}}/3$. Note that the threshold condition is in this case equivalent to demanding that $\breve{f}_{\mathrm{CSF}} < \tau/3$. To understand the choice of $\mathbf{L}_{\mathrm{CSF}} = 1/\sqrt{D}$ as an implementation of the assumption that f_{CSF} is zero in the fit, it is again sufficient to consider a scenario where the CSF fraction is spread out over D directions, and then apply Eq. (4.13) to obtain the regularization term $\|\mathbf{Lf}\|_2^2$ for the total CSF fraction: it would in that case evaluate to $D \cdot (f_{\mathrm{CSF}}/D)^2$. On the other hand, applying Eq. (4.14) directly to the total CSF fraction leads to $(f_{\mathrm{CSF}}/\sqrt{D})^2$. However, since both of these expressions are equal (they both reduce to $(f_{\mathrm{CSF}})^2/D$), a good balance has been established between Eq. (4.13) and Eq. (4.14). Both the definition and the explanations for $\mathbf{L}_{\mathrm{GM}}$ are fully analogous to $\mathbf{L}_{\mathrm{CSF}}$.

The regularization matrix $\mathbf{K}$ also has an overall layout

$$\mathbf{K} = \begin{bmatrix} \mathbf{K}_{\mathrm{CSF}} & 0 & 0 \\ 0 & \mathbf{K}_{\mathrm{GM}} & 0 \\ 0 & 0 & \mathbf{K}_{\mathrm{WM}} \end{bmatrix} \tag{4.15}$$

but only acts on the WM fractions; hence $\mathbf{K}_{\mathrm{CSF}} = 0$ and $\mathbf{K}_{\mathrm{GM}} = 0$. The remaining $\mathbf{K}_{\mathrm{WM}}$ is a $D \times D$ Laplacian matrix corresponding to the fixed large set of uniformly distributed directions z_d. We define it as

$$\mathbf{K}_{\mathrm{WM},i,j} = \begin{cases} \#\mathrm{neighbors}\,(z_i) & \text{if } i = j \\ -1 & \text{if } z_i \text{ and } z_j \text{ are neighbors} \\ 0 & \text{otherwise} \end{cases} \tag{4.16}$$

where $\#\mathrm{neighbors}\,(\mathrm{z_i})$ denotes the number of neighbors of direction z_i and the actual neighbor relation between directions results from a Delaunay triangulation

of $2 \cdot D$ points that represent the D directions on the unit sphere, as shown in Fig. 4.1. Such a Delaunay triangulation can be obtained by computing the convex hull of the points in the 3-dimensional Euclidean space.

To intuitively understand how adding both regularization terms resolves the potentially underdetermined nature of the system in Eq. (4.9), we can reformulate the minimization problem in Eq. (4.10) as

$$\hat{\mathbf{f}} = \min_{\mathbf{f}} \left\{ \left\| \begin{bmatrix} \mathbf{R} \\ \lambda\mathbf{L} \\ \kappa\mathbf{K} \end{bmatrix} \mathbf{f} - \begin{bmatrix} \mathbf{s} \\ 0 \\ 0 \end{bmatrix} \right\|_2^2 \right\} \tag{4.17}$$

which leads to the observation that each independent non-zero row of either regularization matrix adds another equation to the system. The reason the original system could be underdetermined lies in the use of the large number D of directions for which a f^d_{WM} is to be estimated. However, the regularization matrix $\mathbf{K}$, or more specifically the $D \times D$ Laplacian matrix $\mathbf{K}_{\mathrm{WM}}$, has this problem – mathematically speaking – covered. In principle (and ignoring the CSF and GM fractions), only a single measurement is required to obtain a unique solution for the complete set of f^d_{WM} fractions. In reality though, the quality of the outcome is of course still dependent on the angular resolution of the measured data; the less data available, the more the outcome relies on the assumptions of the regularization priors. From the second iteration and onwards, the regularization matrix $\mathbf{L}$ also comes into play, adding extra equations to the system. Once typical (large) angular regions of f^d_{WM} fractions that are assumed to be zero start to appear, the equations added by non-zero rows in $\mathbf{L}_{\mathrm{WM}}$ will encompass some of the prior information already provided by the corresponding rows in $\mathbf{K}_{\mathrm{WM}}$; but also make it stronger. Basically, the assumption of a certain f^d_{WM} and all neighboring WM fractions being zero, fully implies the assumption of a zero-valued Laplacian for the corresponding direction. Both regularization priors reinforce each other to optimally result in coherent and constant zero-valued f^d_{WM} angular regions where necessary. Also, the remaining number of *non-zero* f^d_{WM} should typically be much lower than the initial D; this consequently reduces the underdetermined nature of the initial problem even further. While $\mathbf{L}_{\mathrm{WM}}$ plays a big role in establishing the angular location and extent of these non-zero regions, it does not act within the regions themselves: here, the combination of the data term and $\mathbf{K}_{\mathrm{WM}}$ ensure a good and coherent fit. The unique interactions between the three terms of Eq. (4.10) should thus result in a sparse *presence* of populations, without enforcing a sparse *appearance* of any single population *per se*.

Finally, to recover f_{CSF} and f_{GM}, and distinguish them successfully from the WM fraction, we need more than a single b-value. Just distinguishing f_{CSF} and f_{GM} from each other clearly requires two b-values, as the angular information

within the shells is useless to differentiate two IVFs. Distinguishing both of them from the WM fraction theoretically requires yet another extra b-value. One could of course also reason that the typically expected anisotropic nature of the WM itself (and the fact that the f^d_{WM} fractions are the only mechanism to allow for anisotropy in our MSMT representation) provides sufficient information to differentiate it from the IVFs. However, that again implies an assumption which could adversely influence the outcome in case of highly dispersed (or almost isotropic) WM. Additionally, noise in CSF or GM regions might too easily be regarded as WM under such assumptions. To simply avoid these assumptions altogether, 3 b-values are required. Note that $b = 0$ measurements also count towards this total; *i.e.* the requirement ends up being what many would *informally* refer to as "2 shells of DWIs and a number of $b = 0$ images".

4.2.3 A novel directionally-encoded color (DEC) WM map

A good way to visualize the fractions of our MSMT representation, is to present maps of the normalized triplet $(\breve{f}_{\mathrm{CSF}}, \breve{f}_{\mathrm{GM}}, \breve{f}_{\mathrm{WM}})$, either separately or possibly combined by use of the three RGB color channels. This does, however, show no information on the separate $\breve{f}^d_{\mathrm{WM}}$. To efficiently and visually assess directional information, a common practice is to use directionally-encoded color (DEC) maps, *e.g.* the trademark DEC FA map that DTI is well known for (Pajevic and Pierpaoli, 1999). Therefore, we propose a novel DEC WM map which maximally exploits all information in the complete set of $\breve{f}^d_{\mathrm{WM}}$ to obtain a single representative DEC value, which is subsequently weighted by $\breve{f}_{\mathrm{WM}}$ itself. It is specifically designed to address a multitude of shortcomings of the DEC FA map that are known to frequently cause over- or downright mis-interpretation (Jones et al., 2013).

We define our DEC WM map's values as

$$\mathrm{DEC}_{\mathrm{WM}} = \frac{\sum_{d=1}^{D} \breve{f}^d_{\mathrm{WM}} \cdot \mathrm{DEC}^{z_d}}{\left\| \sum_{d=1}^{D} \breve{f}^d_{\mathrm{WM}} \cdot \mathrm{DEC}^{z_d} \right\|_2} \cdot \breve{f}_{\mathrm{WM}} \tag{4.18}$$

where DEC^{z_d} denotes a 3-element vector of RGB values that is deemed appropriate to represent direction z_d. We employ the convention that is typically used to represent directional information in the domain of DWI; *e.g.* red for mediolateral, green for anteroposterior, and blue for superoinferior (Pajevic and Pierpaoli, 1999). In general, we simply obtain DEC^{z_d} from the elementwise absolute value of a 3-element unit vector, expressed in Cartesian coordinates, along z_d. The resulting DEC^{z_d} scheme for the set of $D = 300$ directions used in this work, is presented in Fig. 4.1. Note that the normalization (*i.e.* the denominator $\| \dots \|_2$) and subsequent reweighting by

$\breve{f}_{\mathrm{WM}}$ in Eq. (4.18) are still essential when each DEC^{z_d} is a unit vector; as we have $\| \sum_{d=1}^{D} \breve{f}_{\mathrm{WM}}^{d} \cdot \mathrm{DEC}^{z_d} \|_2 \leq \breve{f}_{\mathrm{WM}}$ in this case.

Overall, there are two major differences between the proposed DEC WM map and the classical DEC FA map. First, the DEC of the DEC WM map is calculated in a much more robust way that allows each $\breve{f}_{\mathrm{WM}}^{d}$ to have a fair and directly proportional contribution to the final established DEC. In regions of multiple crossing WM populations, the individual populations' colors are intuitively (and again proportionally) mixed. In contrast, the DEC of the DEC FA map directly depends on the direction of the diffusion tensor's principal eigenvector; which is known to be ill-defined in voxels with multiple crossing (or highly dispersed) WM populations, as well as in cases of partial voluming with either CSF or GM. In such voxels, it tends to be oriented *almost* randomly along the direction of either of the populations, or even none of them at all; as the measurement noise could just as well make the difference. Second, the $\breve{f}_{\mathrm{WM}}$ weighting directly represents the relative presence of WM, while FA is often erroneously seen as a measure of "WM integrity". The latter again is a source of misinterpretation for voxels containing crossing or dispersed WM populations, as the FA will show remarkable drops in such regions (Jones et al., 2013). The total relative $\breve{f}_{\mathrm{WM}}$, on the contrary, simply equates to the sum of all individual relative $\breve{f}_{\mathrm{WM}}^{d}$; and is thus much more meaningful and easier to interpret.

4.2.4 Preservation of Principal Volume Fractions (PPVF) retransformation

Once the MSMT representation has been fitted to the measured signal, applying preservation of principal volume fractions (PPVF) retransformation is quite straightforward. Given an affine transformation matrix for each voxel (either directly resulting from a global affine image transformation, or provided by the Jacobian matrix each voxel in case of a non-rigid deformation field), the forward transformation is applied directly to the set of D directions that are used for the MSMT representation of the measured q-space signals. Using the individually reoriented directions, the model in Eq. (4.1) can be evaluated again in a forward sense for *any* given set of gradient directions (and those b-values for which the response functions allow evaluation), *e.g.* by use of the matrix notations in Eq. (4.8) and Eq. (4.9). The magnitudes of the original fractions f_{CSF}, f_{GM} and each $f_{\mathrm{WM}}^{z_d}$ are left unchanged throughout this operation.

To initially fit the MSMT representation, the matrix $\mathbf{R}$ only has to be calculated once for the whole image. When performing *non-rigid* deformation, on the other hand, we end up with a different set of D directions in each voxel; thus each yielding a unique $\mathbf{R}$. Depending on how the WM response function itself

is represented, it might be computationally expensive to calculate (*i.e.* reorient) and/or evaluate each new $R^{z_d}_{\mathrm{WM}}(u,b)$ in each single voxel. However, since the WM response function $R_{\mathrm{WM}}(\theta,b)$ features axial symmetry around the $\theta = 0$ direction, we can exploit the nifty property

$$R^{z}_{\mathrm{WM}}(u,b) = R_{\mathrm{WM}}(z\angle u, b) \tag{4.19}$$

where the $\angle$ operator signifies the angle between two directions. We thus only need to be able to evaluate $R_{\mathrm{WM}}(\theta,b)$ quickly. Should the latter still constitute a computationally expensive operation, it is simply possible to pre-evaluate each shell (in a multi-shell setting) of $R_{\mathrm{WM}}(\theta,b)$ for a decent amount of θ-values. The final evaluation of $R_{\mathrm{WM}}(\theta,b)$ for any θ can then quickly be retrieved through interpolation in the 1-dimensional θ-domain. This technique even allows for a "free-form" single fiber WM response function, *i.e.* removing the need for any analytic closed-form model or representation.

We thus refer to our method as PPVF retransformation, because it is a natural generalization of the PPD reorientation strategy used for the diffusion tensor, to an arbitrary representation that contains multiple (tissue) volume fractions. Each separate fraction is reoriented independently while maintaining its own shape and magnitude; and preserving its main orientation throughout the transformation. Because our MSMT representation provides separate IVFs for CSF and GM, anisotropy can not "magically" be created in these tissues' regions during deformation of a dataset. The WM, however, is captured by an AVF spread over D directions. This allows the distribution of WM to naturally deform; *e.g.* by allowing changes in angles between crossing WM populations, or even changes in dispersion of individual populations. Hence, it should be clear that some measures' values, such as the FA, calculated from the data before and after PPVF retransformation are able to change. While *e.g.* the FA is rotationally invariant, it is *not* "transformationally" invariant. Some transformational invariant measures are provided by the MSMT representation itself: f_{CSF}, f_{GM} and f_{WM}; as well as their normalized counterparts $\breve{f}_{\mathrm{CSF}}$, $\breve{f}_{\mathrm{GM}}$ and $\breve{f}_{\mathrm{WM}}$. The total (or average) amount of q-space signal for each shell also remains transformationally invariant.

4.2.5 Multi-subject multi-channel registration

We developed a new multi-subject multi-channel diffeomorphic matching algorithm (Dhollander et al., 2011c) in order to match image data of a group of subjects simultaneously in their average space. The algorithm directly compares all the subjects' image volumes among themselves throughout the registration process, *i.e.* the deformations of the image volumes are never guided by an intermediate (fuzzy) template. Upon convergence, however, a template can of

course still be obtained by averaging out the deformed image volumes. The algorithm itself employs the well known and popular demons forces (Thirion, 1998). We chose the symmetric variant of the forces that takes into account the gradients of both images. It was shown that this has theoretical as well as practical advantages (Vercauteren et al., 2009). The force $\vec{V}(I, J)$ between 2 images I and J is defined as:

$$\vec{V}(I,J) = -\frac{(I-J)\left(\vec{\nabla}I + \vec{\nabla}J\right)/2}{\left\|\left(\vec{\nabla}I + \vec{\nabla}J\right)/2\right\|_2^2 + (I-J)^2/(2\epsilon)^2} \tag{4.20}$$

where ϵ acts as a soft maximum on $\|\vec{V}(I,J)\|_2$. The force is symmetric since $\vec{V}(I,J) = -\vec{V}(J,I)$. The algorithm can be provided with data of N subjects, containing possibly M channels of information each. At each iteration, given all currently deformed subjects, *every* subject attempts to move/deform closer to *all* other subjects. To achieve this, an unconstrained correspondence update field $\vec{U}(I_i)$ is calculated for each subject I_i by

$$\vec{U}(I_i) = \sum_{\substack{1\le k\le M \\ 1\le j\le N \\ j\neq i}} \frac{\vec{V}(I_{i,k}, I_{j,k})}{(N-1)M} \tag{4.21}$$

where $I_{i,k}$ represents a channel of this subject I_i, while $I_{j,k}$ refers to the corresponding channel of another subject I_j. Hence, Eq. (4.21) boils down to the average force that all $N-1$ other subjects exert upon this subject, over all M channels. The algorithm proceeds as diffeomorphic demons (Vercauteren et al., 2009), but for every separate subject simultaneously. First, fluid regularization is imposed by smoothing each $\vec{U}(I_i)$ with a Gaussian filter. Next, the diffeomorphic update step is performed by composing the current deformation fields with the fast vector field exponentials of the correspondence update fields. Finally, elastic regularization is imposed by smoothing each resulting updated deformation field with a Gaussian filter. All subjects' deformed channels can now be obtained according to the new total deformation fields by interpolating the original data. These can then serve again as the input to Eq. (4.21) for the subsequent iteration.

The only parameters to be set are ϵ and the standard deviations of both Gaussian regularization filters. We always set all three of these to the same value, and then gradually lower this value in a few discrete steps to result in a careful multi-scale approach. At each "scale", the algorithm is allowed to converge before moving on to the next one. We employ values of 3, 2, 1 and finally 0.5 (voxel units). In our experience, this typically results in a robust registration process. Finally, for the choice of channels, we employ the triplet of our newly

obtained normalized tissue fractions ($\breve{f}_{\mathrm{CSF}}, \breve{f}_{\mathrm{GM}}, \breve{f}_{\mathrm{WM}}$), as these should feature comparable (normalized) magnitudes and are also transformational invariants of the PPVF retransformation that we will apply afterwards.

4.2.6 Data and preprocessing

For the purpose of illustrating our proposed methods, we use open access data from the Q3 data release of the Human Connectome Project (Van Essen et al., 2013). These data include high quality multi-shell DWI acquisitions; details on scanner hardware, acquisition setup and image reconstruction algorithm can be found in Ugurbil et al. (2013) and Sotiropoulos et al. (2013). The data was furthermore preprocessed by a state-of-the-art pipeline that includes susceptibility induced EPI distortion correction using reverse phase encoded images (Andersson et al., 2003), combined eddy-current induced field inhomogeneity and head motion correction with appropriate reorientation of the gradients (Andersson et al., 2012), *rigid* transformation to a standard space (*no* deformation or scaling) and brain masking. Most of the processing steps are performed using the publicly available *FSL* (Jenkinson et al., 2012) and *FreeSurfer* (Fischl, 2012) packages. More details on the complete preprocessing pipeline are provided in Glasser et al. (2013).

The resulting preprocessed DWI datasets all feature a spatial resolution of $1.25 \times 1.25 \times 1.25\,\mathrm{mm}^3$. They include 18 non-DWI volumes ($b = 0$), and 3 shells ($b = 1000\,\mathrm{s/mm}^2$, $b = 2000\,\mathrm{s/mm}^2$ and $b = 3000\,\mathrm{s/mm}^2$) each consisting of DWI measurements for their own *unique* set of 90 gradient directions. The complete set of 3×90 gradient directions contains 270 *unique* directions, that are optimized to be distributed uniformly for each separate shell as well as for the complete set; using a variant of the electrostatic repulsion algorithm specifically designed for multiple shells (Caruyer et al., 2013).

For this work, we randomly selected a group of 14 subjects, not guided by any specific criteria. This group may (or may not) even include twins. Using *MRtrix* (Tournier et al., 2012), we then spatially regridded these data ourselves to a resolution of $2 \times 2 \times 2\,\mathrm{mm}^3$. This final step was performed intentionally to obtain a more *realistic* spatial resolution (as compared to current commonly achievable resolutions). This should render the challenge of obtaining tissue response functions directly from the data much harder, especially for the GM. Note that we regard the data from this point on as simply having a spatial resolution of $2 \times 2 \times 2\,\mathrm{mm}^3$: we will for no purpose whatsoever access any information from the original data that existed before this regridding step.

Finally, we will also use the accompanying T2 weighted volumes and binarized *FreeSurfer* cortical ribbons of these 14 subjects (both at a spatial resolution of

$0.7 \times 0.7 \times 0.7\,\text{mm}^3$). These will assist us in a few *ad hoc* strategies to obtain voxels for measuring the response functions. The binarized cortical ribbons feature a label for the region between the inner and outer cortical surface (*i.e.* the cortex), as well as for the region "within" the inner cortical surface (*i.e.* the subcortical region). Note that the latter is *not* simply a WM segmentation: it includes the lateral ventricles and subcortical deep GM structures, and excludes non-cerebral white matter (*i.e.* in the brain stem and the cerebellum).

4.2.7 Experiments

The experiments for which results are presented in this chapter, are designed to give the reader a general feeling for (some of) the properties of our proposed MSMT representation, deliver a proof of concept of our fitting strategy in a realistic setting while providing some extra insight in the general behaviour of the regularization parameters, show the MSMT representation's applicability in the context of PPVF retransformation, and finally apply this technique to obtain an average MSMT template containing average multi-shell q-space signals. While this already provides a decent amount of novel information to process, we want to stress our awareness of the fact that this still leaves a lot to be investigated: the MSMT concept provides a highly flexible framework that is open to changes, improvements and adaptions for specific purposes. Some visionary brainstorming on such possibilities is provided in the final discussion.

A first requirement for fitting our proposed MSMT representation consists of obtaining CSF, GM and single fiber WM response functions. Existing techniques, such as constrained spherical deconvolution (CSD) (Tournier et al., 2004, 2007), also require the estimation of a single fiber WM response function. Such a response function is typically obtained from the data itself, and this is most commonly achieved using rather *ad hoc* approaches, *e.g.* including a simple threshold on FA value to find voxels that are assumed to contain a single fiber WM configuration. In this work, we also employed three very simple and equally *ad hoc* approaches to obtaining response functions for each tissue class. While much more advanced methods can certainly be devised, we simply aim to illustrate that it is still not very hard to obtain *reasonable* response functions (even for DWI data at a spatial resolution of $2 \times 2 \times 2\,\text{mm}^3$) that subsequently perform well in terms of a *decent* MSMT-fitted outcome. These are the procedures specifically applied in this work:

- CSF response function: the $0.7 \times 0.7 \times 0.7\,\text{mm}^3$ T2 weighted volume was thresholded (at a value of 500) to obtain a region assumed to contain CSF; the binary mask was regridded to the $2 \times 2 \times 2\,\text{mm}^3$ space of the

DWI data; in the resulting voxels, the average (non-)DWI signal for each shell was obtained to represent the CSF response function.

- GM response function: the $0.7 \times 0.7 \times 0.7\,\mathrm{mm}^3$ binarized cortex region from the *FreeSurfer* cortical ribbon was eroded (2 passes) to obtain a region assumed to contain GM; the binary mask was regridded to the $2 \times 2 \times 2\,\mathrm{mm}^3$ space of the DWI data; in the resulting voxels, the average (non-)DWI signal for each shell was obtained to represent the GM response function.
- Single fiber WM response function: the $0.7 \times 0.7 \times 0.7\,\mathrm{mm}^3$ binarized subcortical region from the *FreeSurfer* cortical ribbon was regridded to the $2 \times 2 \times 2\,\mathrm{mm}^3$ space of the DWI data; a further FA threshold (at a value of 0.7) retained voxels assumed to contain single fiber WM; in these voxels, the average non-DWI signal ($b = 0$) was obtained to represent the single fiber WM $b = 0$ response; furthermore, the DWI signals ($b \neq 0$) for each shell were reoriented by aligning the principal eigenvectors of the diffusion tensors; average spherical harmonics (SH) fits of these signals (up to SH order 8, only using even SH orders, and at zero phase factor for axial symmetry) represent the single fiber WM response for each of these shells.

The latter strategy to obtain a single fiber WM response function, is mostly similar to the one employed in Tournier et al. (2004) and Tournier et al. (2007). It was recently shown that an SH order of 8 is sufficient to capture all angular features of the DWI signal under realistic acquisition conditions and up to $b = 5000\,\mathrm{s/mm}^2$ (Tournier et al., 2013a). Note furthermore that the use of SH basis functions to *measure* the response function has no actual impact on the rest of the MSMT fitting procedure; it only limits the angular frequency content of the response function itself. All of the above *ad hoc* pipelines were fully automated using *MRtrix* (Tournier et al., 2012) to obtain the three response functions for each of the 14 subjects. These results were finally averaged over all subjects to obtain a unique response function for each tissue class.

In a next step, we employed the response functions to fit the MSMT representation to the data of a single subject. We explored different (combinations of) values for the regularizaton parameters λ and κ and investigated their impact on the quality of the resulting fit, as well as on the number of iterations that is required for convergence. The fitting procedure for each voxel is always allowed an absolute maximum of 11 iterations: 1 initial iteration (where no fractions are assumed to be zero yet) and 10 subsequent iterations that may adjust their assumptions about the set of zero valued fractions according to the outcome of the previous iteration. Should this set not change between two consecutive iterations, the algorithm has converged.

Based on the outcome of the previous investigation of the regularization parameters, we will decide to continue with a fixed reasonable choice that suits the overall purpose of this work. MSMT maps of the normalized relative tissue fractions and the novel DEC WM map, are compared to maps resulting from a fit of the recently proposed NODDI model as well as the FA and DEC FA maps from the DTI model. Furthermore, we show the impact of the PPVF retransformation on the MSMT generated q-space signal as resulting from an artificial transformation (a simple global shear).

The MSMT representation was then fitted to all 14 subjects. The triplet of normalized relative tissue fractions was employed to guide a multi-subject multi-channel registration, as described previously. The resulting 14 deformed (and PPVF retransformed) MSMT representations were combined to obtain an average MSMT template. Finally, the MSMT representation of this template was evaluated again to obtain an average multi-shell q-space signal template.

4.3 Results

4.3.1 Response function estimation

The main strategy of the *ad hoc* pipelines for response function estimation from the data always starts by locating voxels that optimally present pure (uncontaminated) examples of the required tissue classes; *i.e.* CSF, GM and single fiber WM. Note that, even though we used information from higher resolution sources (*e.g.* a $0.7 \times 0.7 \times 0.7\,\mathrm{mm}^3$ T2 weighted volume) to assist at this task, the final selected regions consist of binary regions of voxels on the grid of the DWI data itself (*i.e.* at a spatial resolution of $2 \times 2 \times 2\,\mathrm{mm}^3$). The higher resolution information was *not* used to interpolate or weight the DWI data. The regions resulting from our *ad hoc* pipelines are presented in Fig. 4.2 for three slices (axial, coronal and sagittal) of a single subject. The selected CSF voxels are found mostly in the ventricles, but also surrounding the brain and within some larger sulci. The GM voxels resulting from the pipeline lie within the cortex. We intentionally excluded subcortical deep GM structures, as these resemble a mix of white and gray matter (*i.e.* they also contain a relevant amount of myelinated axons). For similar reasons, the cerebellum is intentionally not included: at a spatial resolution of $2 \times 2 \times 2\,\mathrm{mm}^3$, most (if not all) of its cortical voxels will show a mixture of WM and GM. The GM certainly provides the biggest challenge with regard to selecting representative and uncontaminated voxels at this resolution, as human cortical thickness typically averages around $\sim 2.5\,\mathrm{mm}$. However, Fig. 4.2 shows that we could still obtain quite a reasonable amount of them; hence, there is even an extra margin to be more careful and

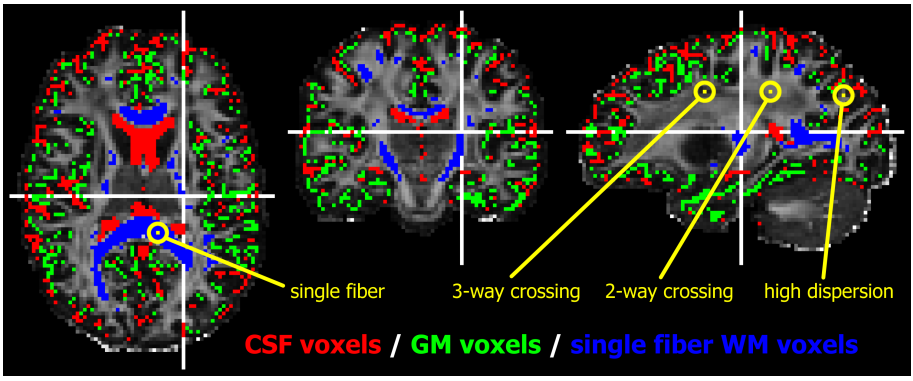

Figure 4.2: The voxels used to estimate the CSF (*red*), GM (*green*) and single fiber WM (*blue*) response functions; as resulting from three *ad hoc* pipelines and visualized for a single subject (the final response functions are averaged over all 14 subjects). The white crosshairs indicate the location of the presented slices relative to each other. The encircled voxels feature particular WM configurations for which MSMT fitted results are investigated. Everything is overlaid on a FA map for spatial reference.

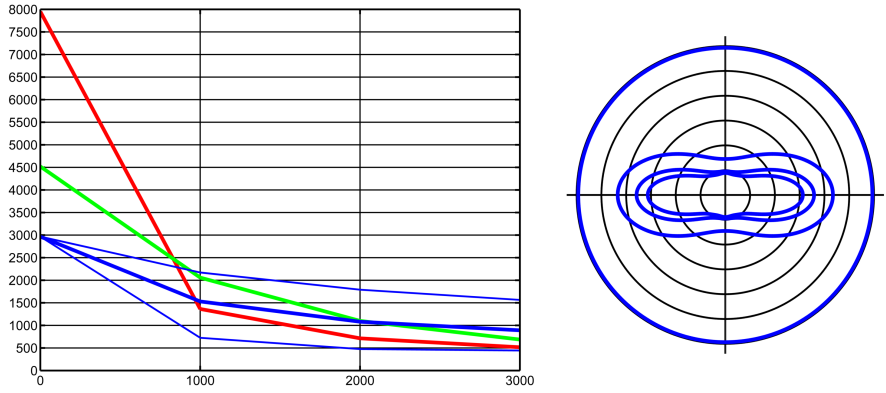

Figure 4.3: The estimated average CSF (*red*), GM (*green*) and single fiber WM (*blue*) response functions. *Left:* the (average) amplitude of the response functions in function of b-value. For the single fiber WM response function (*blue*), the thick line represents the average amplitude over the whole angular domain, while the thin lines indicate the amplitude along the axial (least hindered/restricted) and radial (most hindered/restricted) directions. *Right:* the single fiber WM response function for each b-value, in function of inclination angle θ. The vertical axis represents the central axis of symmetry ($\theta = 0$).

restrictive. Regarding the voxels used to measure the response functions of CSF and GM in this work, it is reassuring that there still exists a small gap between both classes in many regions in Fig. 4.2; as we typically expect a single voxel wide layer of CSF-GM partial volumed voxels between both. Finally, the single fiber WM voxels are mostly determined by the FA threshold. This commonly applied technique is often combined with some strategy to exclude spurious high FA voxels at the edge of the brain mask (resulting from the meninges and other tissues). In this work, we used the subcortical region for this purpose, but an eroded brain mask should be equally suitable.

As described previously, we use the obtained regions to measure the average response functions for each subject, and average the outcomes over all subjects to obtain the final three unique response functions. Plots of these response functions are provided in Fig. 4.3. The signal decays confirm that we have the most free diffusion in CSF, and the most hindrance/restriction of diffusion in WM on average (and specifically even more along the radial direction). The angular plot of the single fiber WM response function also shows an increasing contrast for higher b-values. Based on these plots, we could even come up with *ad hoc* strategies to obtain regions for estimating the response functions, based only on the DWI data itself. For the CSF, we could simply apply a threshold on the (average) $b = 0$ volume. For the GM, it seems feasible to apply a threshold on the *average* $b = 1000\,\mathrm{s/mm^2}$ volume. For the WM, we can stick with the simple FA threshold.

4.3.2 Effect of regularization parameters

We investigated large ranges of (combinations of) both regularization parameters λ and κ. For practical reasons though, we only present the reader with the qualitative results for a limited range of values. These ranges have been specifically selected to optimally illustrate the main behavior of both parameters and their overall effect on the results. Hence, we provide results for each combination of $\lambda = 10^2, 10^4, 10^6, 10^8$ and $\kappa = 10^2, 10^3, 10^4$.

The top row of Fig. 4.4 shows the resulting triplet of *normalized* relative fractions $(\breve{f}_{\mathrm{CSF}}, \breve{f}_{\mathrm{GM}}, \breve{f}_{\mathrm{WM}})$. As the triplet sums to 1, the "normal" range of values equals 0–1 in case no negative values are present. The color map in Fig. 4.4 visualizes this normal range in grayscale, while negative values go to cyan/green and values beyond 1 go to red/yellow. When λ is set (relatively) low, negative fractions are clearly still present and are often accompanied by values beyond 1 for the complementary fraction(s). In this case, overall, the same pattern and logic is seen for each main tissue class: due to overfitting, the purest voxels of each tissue class show an exaggerated contrast by engaging negative or otherwise

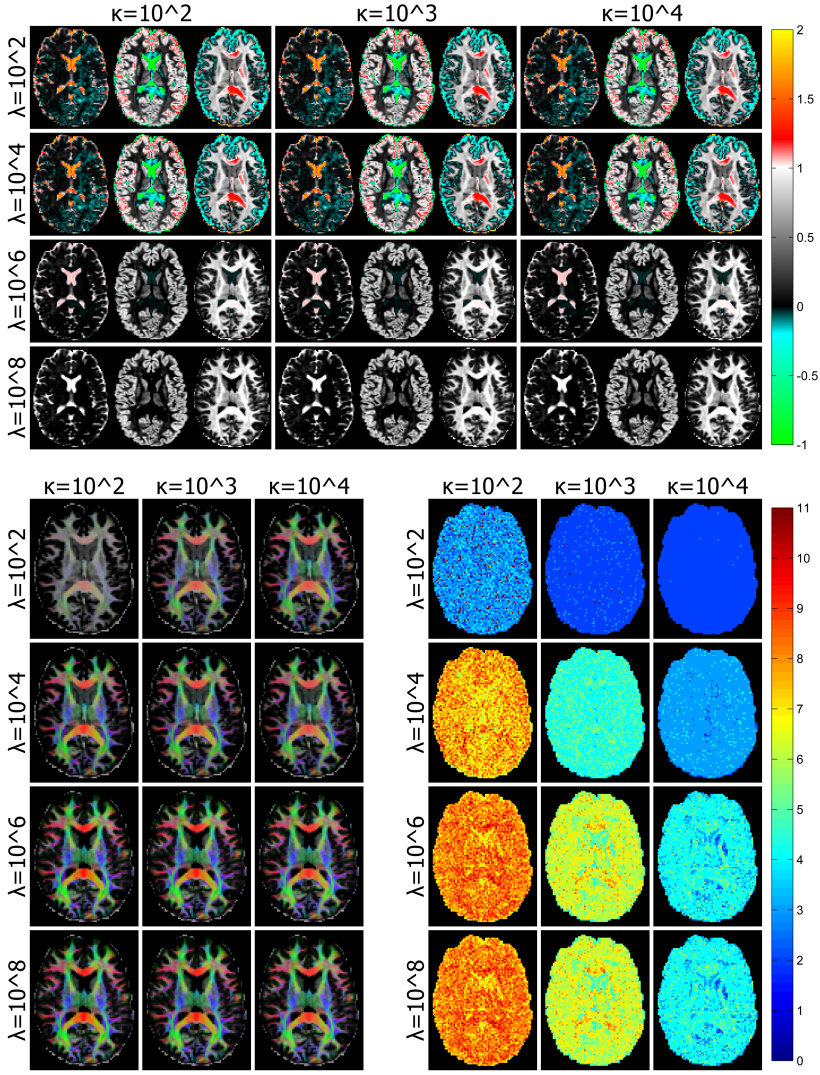

Figure 4.4: Outcomes of the MSMT fitting for different combinations of regularization weights λ and κ. *Top:* each block shows a triplet of *normalized* relative CSF, GM and WM fractions. *Bottom left:* the DEC WM map, where the DEC is obtained only from the non-negative normalized directional WM fractions. The negative directional fractions are taken into account for the weighting by the total normalized WM fraction though. *Bottom right:* the number of iterations needed to fit the MSMT representation in each voxel. This number includes the initial iteration where no fractions are assumed to be zero yet.

"manipulated" fractions of other tissue classes. In large parts of the CSF, we have $\breve{f}_{\mathrm{CSF}} > 1$, combined with $\breve{f}_{\mathrm{GM}} < 0$, but also manipulated values of $\breve{f}_{\mathrm{WM}}$ that are significantly *larger* than 0 when – obviously – they shouldn't be (they even reach comparable magnitudes to those in subcortical deep GM structures). In the purest cortical GM voxels, we have $\breve{f}_{\mathrm{GM}} > 1$, combined mostly with $\breve{f}_{\mathrm{WM}} < 0$, and sometimes slightly negative values of $\breve{f}_{\mathrm{CSF}}$. Note that $\breve{f}_{\mathrm{WM}}$ additionally shows negative values in larger parts of the cortex, hinting at the fact that $\breve{f}_{\mathrm{GM}}$ might also be overestimated in these regions (even though it isn't valued beyond 1). In the purest single fiber WM regions, we have $\breve{f}_{\mathrm{WM}} > 1$, combined mostly with $\breve{f}_{\mathrm{GM}} < 0$, but also $\breve{f}_{\mathrm{CSF}} < 0$ to a lesser extent. These overall findings for (relatively) low values of λ possibly suggest that our estimated response functions are slightly lacking the maximum contrast needed to fully differentiate between the tissue classes in an (almost) completely data-driven manner. While a more careful/restrictive approach to measuring the response functions might have already yielded better "constrained" outcomes out of the box at this stage, raising the value of λ clearly renders the fitting procedure itself very robust against the observed side effects. At $\lambda = 10^6$, most values are already within the 0–1 range; only the CSF still shows a very slight resistance. Looking back at Fig. 4.2, we notice that mostly the region in which we estimated the CSF response function might have been a bit too large; potentially including some partial volumed voxels at the CSF-WM boundary and as such slightly reducing the contrast of the measured CSF decay. If we set $\lambda = 10^8$, however, negative (and *very* small positive) fractions, as well as fractions beyond 1, are virtually absent; *i.e.* they are reduced to an order of magnitude that is insignificant relative to the actual estimated fractions. Other anomalies, such as significant non-zero values for $\breve{f}_{\mathrm{WM}}$ in the CSF, are also eliminated. Raising the value of λ to even higher orders of magnitude (*results not shown*) doesn't improve the outcome much further, but it also doesn't degrade (overregularize) its quality; the regularization term simply evolves towards a virtual constraint. Finally, the value of κ does not seem to have any noticeable impact on the $(\breve{f}_{\mathrm{CSF}}, \breve{f}_{\mathrm{GM}}, \breve{f}_{\mathrm{WM}})$ triplet. This confirms that the information to separate the tissue classes, is fully present in the contrast between the average magnitudes of the signal over the different q-space shells; the outcome is not influenced by our fitting procedure's way of dealing with the angular information.

The bottom left of Fig. 4.4 shows the resulting DEC WM maps for each combination of λ and κ. For some of the results, we have to deal with negative $\breve{f}^{d}_{\mathrm{WM}}$ values in the definition of the DEC WM map, Eq. (4.18). The maps presented in Fig. 4.4 ignore negative values of $\breve{f}^{d}_{\mathrm{WM}}$ in the calculation of the DEC, but do take them into account for the final $\breve{f}_{\mathrm{WM}}$ intensity weighting. The results show that, for a combination of low values of both λ and κ, not much color contrast is provided. Providing better regularization by raising either

of both parameters enhances the color contrast. Specifically raising the value of λ also eliminates the presence of $\breve{f}_{\mathrm{WM}}$ in the CSF, and allows for slightly deeper penetration of the color contrast into the inner boundary of the cortex by eliminating the tendency of $\breve{f}_{\mathrm{WM}}$ towards negative values in such regions. Both of these latter observations are directly in line with the aforementioned findings in the previous paragraph. Finally, it is interesting to note that, as we increase the value of λ, the impact of κ on the DEC WM map seems to decrease. Do note that κ certainly still has an influence and its value should *e.g.* not be raised to arbitrarily high values; as overregularization by the accompanying term *will* smooth out the angular (and thus color) contrast at some point. It is, however, nice to see that the regularization term controlled by λ provides an extra source of robustness – also in the angular domain.

The bottom right of Fig. 4.4 shows the number of iterations required by the fitting procedure in each voxel, before convergence was detected (or the absolute maximum of $1 + 10$ iterations was reached). As λ increases, so does the number of required iterations, on average. This makes sense: for lower values of λ, the regularization term lacks the impact required to influence the outcome in a way that allows for an update to the set of fractions assumed to be zero; the algorithm thus converges prematurely. However, when both λ and κ are (too) low, the pattern shows high variability between voxels: some converge prematurely in line with the previous reasoning, others don't seem to converge at all (or at least not within the maximum number of iterations). This consequently represents a rather unstable setting. Raising the value of κ lowers the amount of required iterations, on average. This also makes sense: as the accompanying regularization term reinforces a coherent pattern in the angular domain, it features a unique synergy with the other regularization term (controlled by λ) that results in more coherent sets of fractions being (implicitely) assumed zero. Because we typically also expect such a final pattern, the algorithm is able to converge faster towards such a setting. Finally, as λ is increased towards (relatively) high values, the behavior again stabilizes. Increasing λ further might draw the fractions that are assumed to be zero, even closer to zero; but this has no additional impact on the convergence rate of the fitting procedure that essentially tries to simply establish the set of fractions which are assumed to be zero.

Because the DEC WM maps only present an averaged view in each voxel on the angular configuration of the set of $\breve{f}^{d}_{\mathrm{WM}}$ values, we also specifically investigated several voxels' complete outcome. We provide results for a limited – but representative – set of voxels; whose locations are indicated in Fig. 4.2. They contain specific WM configurations: a single fiber population (actually located in a region where we measured the WM single fiber response function itself), a 3-way crossing, a 2-way crossing and a highly dispersing population.

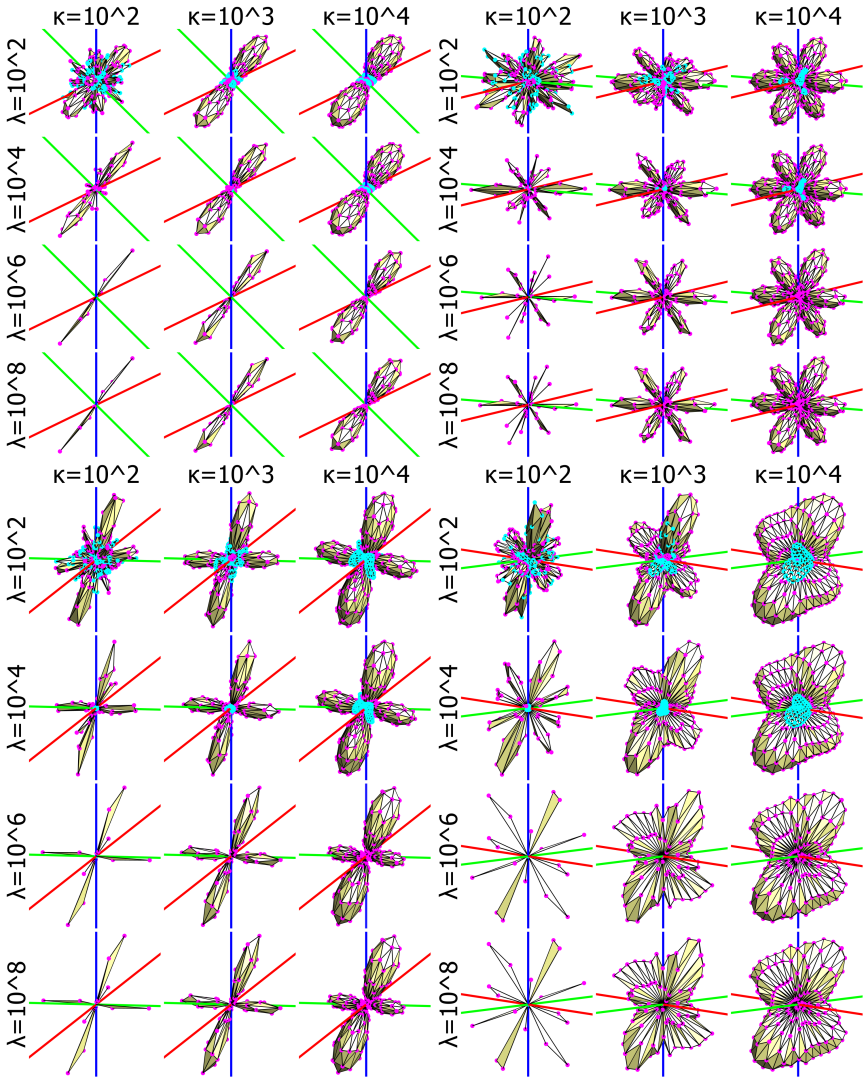

Figure 4.5: Outcomes of the MSMT fitting for different combinations of regularization weights λ and κ, visualizing the directional WM fractions for the encircled voxels indicated in Fig. 4.2. Positive fractions are shown as magenta dots, negative ones as cyan dots. The lines connecting the dots indicate "neighboring" directions, corresponding to the triangulation in Fig. 4.1. Four different WM configurations are presented: single fiber (*top left*), 3-way crossing (*top right*), 2-way crossing (*bottom left*) and high dispersion (*bottom right*). Each result is individually rescaled for optimal visual presentation and comparison.

The outcome for these voxels is provided in Fig. 4.5, again for the range of λ and κ values. Each fraction f^d_{WM} is represented by a dot, and these dots are connected by the same lines that constitute the triangulation in Fig. 4.1. As such, we can easily see whether certain individual fractions relate to neighboring directions or not. Positive fractions are shown in magenta, negative ones in cyan. Some of our earlier findings are immediately confirmed. When both λ and κ are (too) low, the procedure provides a very unstable outcome, featuring large positive as well as negative spurious fractions. Increasing λ to higher values eliminates these spurious fractions. Again, the impact of λ seems to stabilize towards higher values. Increasing κ yields a more coherent outcome – even for low values of λ – but may *unnecessarily* smooth the result for higher values of λ; *i.e.* when the regularization term and the iterative mechanism controlled by λ already provide enough information to reduce the underdetermined nature of the original problem to acceptable levels. While it is clear that a high value of λ, has a desirable effect on the robustness and the overall outcome, it is not so easy to decide upon an "optimal" choice for κ. For the single fiber population, a low κ seems appropriate, as it provides the most sparse outcome (which is in line with the fact that this concerns a voxel that was also used to measure the WM single fiber response function). However, when more dispersion is involved, it is clear that the best outcome does not necessarily match the one with the most sparse *appearance*. As discussed before, the regularization term controlled by κ plays an essential role in synergy with the other regularization term to optimize towards a sparse *presence* of tissue populations. For a low value of κ, we obtain many "angularly disconnected" (*i.e.* non-neigboring) fractions for the 3-way crossing and high dispersion WM configurations. A higher value for κ eliminates this issue and effectively results in a sparse *presence* of tissue populations, but may at some point yield an unnecessary amount of smoothing. Finally, note that a (very small) region of dispersion may always make some sense, even for the single fiber population, due to uncertainty in the data (*e.g.* noise and a limited angular resolution) as well as an unavoidable (slight) mismatch between the measured and "actual" response functions.

Based on our investigations and the presented results in Fig. 4.4 and Fig. 4.5, we decide to stick with $\lambda = 10^8$ and $\kappa = 10^3$ for the remainder of this work. It is clear that a high value of λ is desirable (but beyond the presented range, no significant further impact was observed). For κ, our choice equals a trade-off that provides a minimum level of waranty to obtain the desired sparse *presence* of tissue populations, while not unnecessarily smoothing the results beyond that point. Note that the amount of non-zero fractions obtained in this scenario, still presents a *very* small subset of the total amount of available fractions. As observed in Fig. 4.4, our choice of parameters yields a procedure that converges in 4–7 fast iterations for most voxels, and 8–9 for some more demanding ones. Finally, the reader should be warned that these specifically chosen values for λ

and κ will not automatically apply for their own acquisition setup; *i.e.* they are defined relative to the order of magnitude of the signals, number and choice of shells and gradient directions, etc... The results presented here are merely meant to provide an insight into the impact of both parameters on the final result. The findings also show that (and how) a good fixed choice can be made that suits the expected range of tissue classes and configurations.

4.3.3 MSMT parameter maps

In Fig. 4.6, a typical set of (parameter) maps resulting from the fitted MSMT representation is provided for three slices (axial, coronal and sagittal; the same slice locations as those in Fig. 4.2) and directly compared to maps resulting from a NODDI and DTI fit of the same data. NODDI (Zhang et al., 2012) is a recently proposed model that also expresses the DWI signal as a discrete sum of compartments: an isotropic (CSF) volume fraction, an intra-cellular volume fraction and an extra-cellular volume fraction. It should be noted though, that NODDI is defined in a hierarchical manner: the isotropic (CSF) volume fraction is split from the remainder of the signal (both fractions sum to 1), and the remainder itself is further subdivided into the intra- and extra-cellular volume fractions (both again sum to 1). The typical set of maps consists of the intra-cellular volume fraction, which is also referred to as the neurite density (ND); the isotropic (CSF) volume fraction; and an orientation dispersion (OD) index in a 0–1 range. More details can be found in Zhang et al. (2012).

A map that shows the normalized MSMT tissue fractions using RGB color channels, as well as the individual normalized tissue fraction maps are provided in Fig. 4.6. While the former one yields a quick overview on how everything spatially relates to each other, the latter ones allow for an easier assessment of the individual sources of information. The novel DEC WM map is provided as well. For the NODDI model, we show the ND, the isotropic (CSF) volume fraction and the OD. For the DTI model, the FA map as well as the popular DEC FA map are provided. The first and foremost observation is that the tissue fractions of the MSMT representation correlate well with known anatomy. Subcortical deep GM structures (such as the thalamus, caudate nucleus and putamen) appear as a GM-WM mixture. In the sagittal view, a nice extraction of WM and GM can be appreciated for the cerebellum (which is particularly susceptible to partial voluming due to the spatial scale of its cortical features). Both our MSMT representation and the NODDI model provide a CSF fraction; overall, they also correlate well. Furthermore, the other NODDI maps (ND and OD) relate well to known properties (Zhang et al., 2012) of the corresponding MSMT tissues. It is worthwhile to stress that, due to the hierarchical definition of the NODDI model, the ND (and OD) maps present virtually meaningless

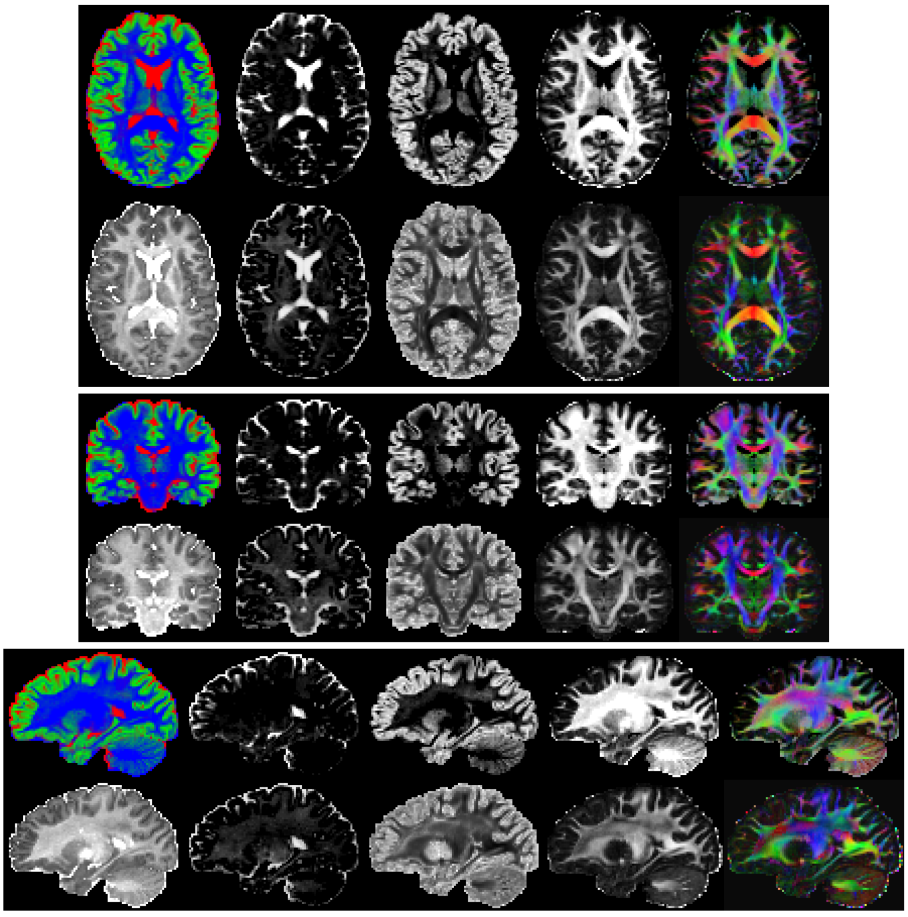

Figure 4.6: Maps obtained from the MSMT representation, compared to NODDI and DTI maps, visualized for the same slice locations (*axial*, *coronal* and *sagittal*) as those in Fig. 4.2. *Top row of each box:* the MSMT maps, from left to right: normalized CSF-GM-WM fractions encoded by RGB; normalized CSF fraction; normalized GM fraction; normalized WM fraction; DEC WM. *Bottom row of each box:* NODDI and DTI maps, from left to right: intra-cellular volume fraction (also called neurite density, ND); isotropic (CSF) volume fraction; orientation dispersion (OD); FA; DEC FA.

values in areas where the isotropic (CSF) volume fraction approaches 1 (*i.e.* the high ND values in the ventricles and sulci don't relate to actual neurites). The ND shows higher values in WM, and lower ones in GM. The ND intensity for the subcortical deep GM structures also lies somewhere in between the WM and GM one. The NODDI model was shown to disentangle key information combined in FA, by ND and OD (Zhang et al., 2012). It was found that FA is *primarily* explained by OD (in most healthy tissues). Hence, FA and OD show an overall (inverse) correlation. This clearly shows why FA is not a specific measure for WM integrity. Consequently, FA drops (and OD is raised) for WM voxels that feature multiple or dispersing populations (*e.g.* in the centrum semiovale, amongst others). The MSMT WM fraction shows no such fluctuations due to complex WM configurations (as does the ND). The FA also drops to almost zero in the entire striatum, rendering it difficult – if not impossible – to differentiate these structures from the nearby CSF in the ventricles; while the MSMT WM fraction still shows the appropriate amount of WM presence. Of course, the DEC FA map suffers from all aforementioned pitfalls in interpretation as well; while the DEC WM doesn't. On top of that, the DEC FA's definition of the DEC adds possible confusion by only being based on the direction of the diffusion tensor's principal eigenvector. This is again an easy source of misinterpretation in voxels containing multiple WM populations. In the sagittal view of Fig. 4.6, for instance, the DEC FA map appears to indicate that the lateral projections of the corpus callosum (red regions) are interupted in a clean-cut fashion by the corticospinal tract (blue region). The DEC WM map, on the other hand, simply suggests that they are crossing (interdigitating) in that particular region by showing a mixed purple color. Similar examples are present in other regions (on the other slices) as well.

Finally, we also provide the reader with a whole-brain coverage of axial MSMT maps in Fig. 4.7 and Fig. 4.8; in order to appreciate and investigate their properties or anatomical validity for different regions. Compared to the DEC FA map, the DEC WM map typically appears to feature a lower contrast. Although this is a simple consequence of the DEC WM map encoding more comprehensive information (*i.e.* WM crossing, dispersion and partial voluming with GM or CSF), several researchers may have gotten used to looking at the vivid colors of the DEC FA map. We therefore also provide a high contrast version of the DEC WM map. This is still the exact same DEC WM map though, only shown at a higher contrast (a smaller intensity window). It might possibly assist in better distinguishing certain WM structures.

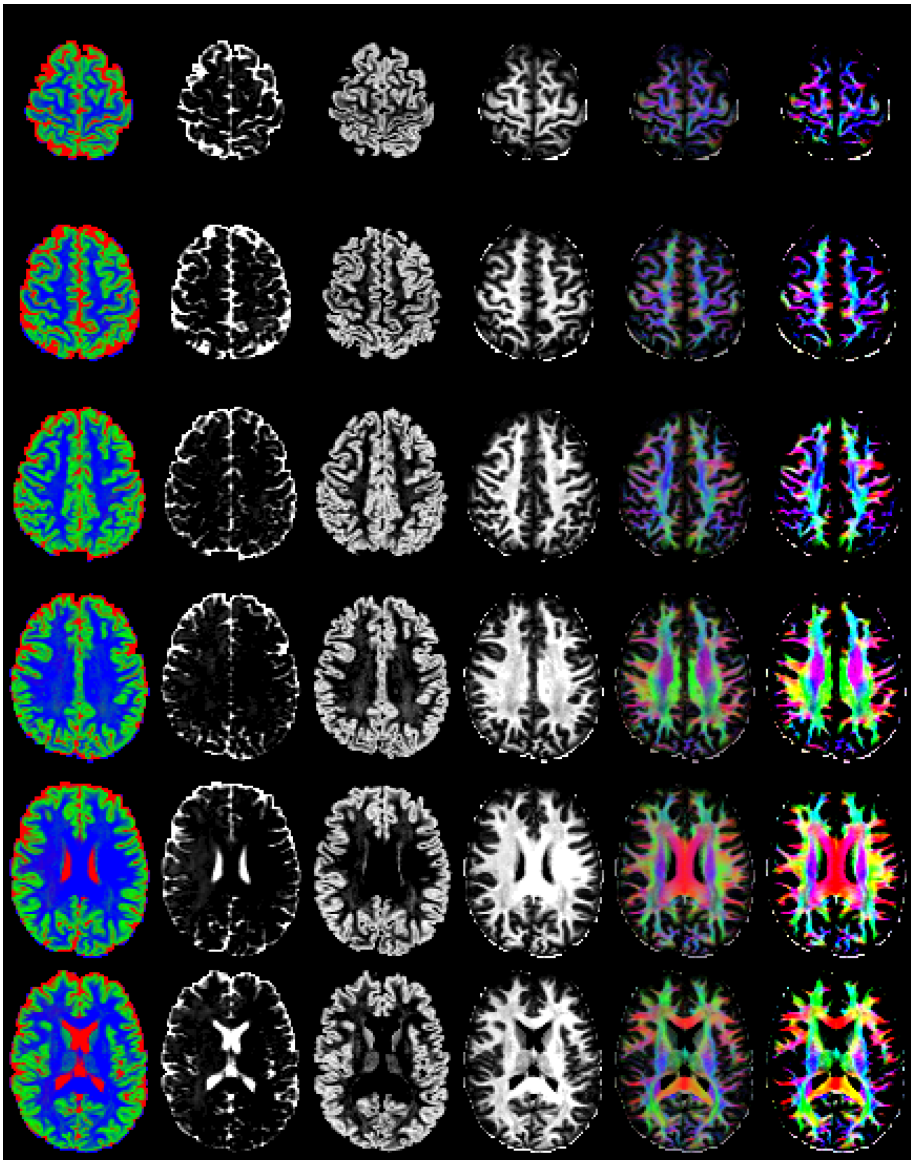

Figure 4.7: Maps obtained from the MSMT representation, visualized at every fourth axial slice for the top half of the brain of a single subject. From left to right: normalized CSF-GM-WM fractions encoded by RGB; normalized CSF fraction; normalized GM fraction; normalized WM fraction; DEC WM; high contrast DEC WM. The latter one is simply the DEC WM shown at a higher contrast, for the sake of enhanced clarity.

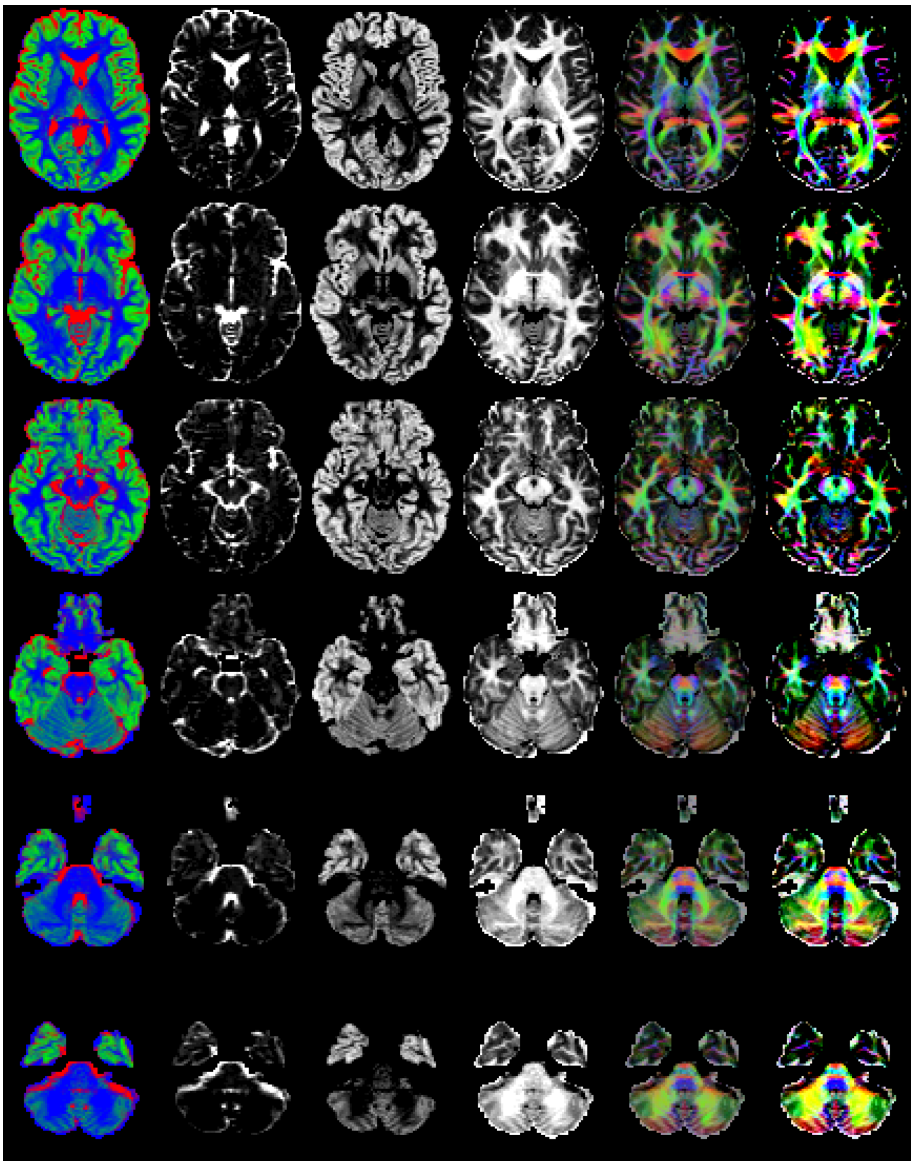

Figure 4.8: Maps obtained from the MSMT representation, visualized at every fourth axial slice for the bottom half of the brain of a single subject. From left to right: normalized CSF-GM-WM fractions encoded by RGB; normalized CSF fraction; normalized GM fraction; normalized WM fraction; DEC WM; high contrast DEC WM. The latter one is simply the DEC WM shown at a higher contrast, for the sake of enhanced clarity.

4.3.4 PPVF retransformation

In order to provide the reader with some insight in the actual effects a PPVF restransformation of the q-space DWI signals entails, we applied an artificial global shearing transformation to the currently obtained MSMT representation. A coronal view on separate steps of the procedure is presented in Fig. 4.9. The "horizontal" shearing should in principle leave the orientation of fiber populations running parallel to an axial plane (*i.e.* horizontally) unaffected. The more a certain fiber populations is running perpendicular to this plane (*i.e.* vertically; parallel to the superoinferior axis), the more it should be affected by this particular shearing.

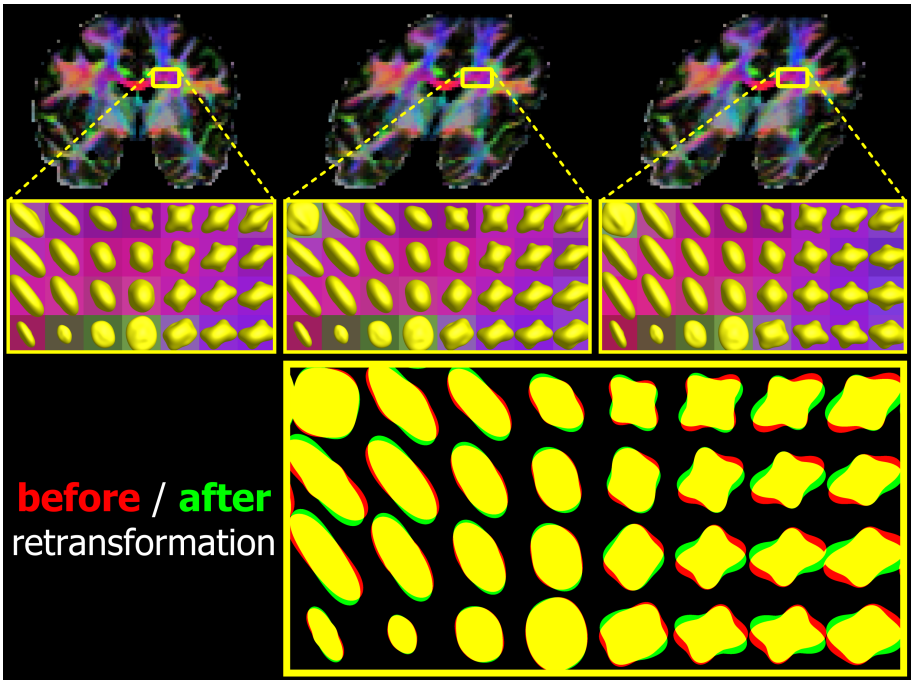

Figure 4.9: PPVF retransformation of q-space signal, illustrated for the $b = 3000\,\mathrm{s/mm^2}$ shell of a single subject. The top row presents a coronal slice before transformation (*left*), after horizontal shearing by spatial interpolation (*middle*) and after subsequent PPVF retransformation of the signal in each voxel (*right*). The DEC WM maps are shown at each stage. The zoomed boxes provide the q-space signal as an overlay. The bottom row allows for an easier comparison between the silhouettes of the q-space signal at the final two stages: before (*red*) and after (*green*) PPVF retransformation.

The results in Fig. 4.9 first present a view on the $b = 3000\,\mathrm{s/mm^2}$ shell before the transformation, as evaluated from the MSMT representation. Next, the transformation is applied spatially by interpolating the MSMT fractions. The PPVF retransformation's task is to correct the now mismatched orientation of the q-space DWI signal related mostly to superoinferiorly running fiber populations. Within the zoomed region of the centrum semiovale, this mostly concerns a population of fibers belonging to the corticospinal tract. The lateral projections of the corpus callosum, however, should be left relatively untouched; since they are running along a more horizontal path. A comparison between the signal before and after retransformation confirms that these requirements are successfully fulfilled by the PPVF strategy. For correct understanding, it is worthwhile clarifying that the q-space signal related to a certain fiber population shows high values along directions *perpendicular* to the population's orientation. Hence, the most affected signals (relating to the corticospinal tract) are found along more horizontal directions. Note that the PPVF retransformation was thus able to change the angle between both populations (the corticospinal tract and the corpus callosum). Furthermore, it will just as well be able to change the amount of dispersion of any population. This implies that the OD, and consequently also the FA, are affected. Consequently, they do not feature transformational invariance. The MSMT fractions, on the other hand, are transformational invariants by definition of the PPVF operation. Finally, the ND should also qualify for transformational invariance, as per its own definition of "intra-cellular volume fraction".

4.3.5 Multi-subject multi-channel registration

Next, we fitted the MSMT representation to the full set of 14 subjects. The resulting normalized relative tissue fractions ($\check{f}_{\mathrm{CSF}}, \check{f}_{\mathrm{GM}}, \check{f}_{\mathrm{WM}}$) are provided in the top box of Fig. 4.10. The maps share similar qualities for all subjects and they are consistent with known normal anatomy. Note that all 14 MSMT representations were obtained in function of the same average response functions; *i.e.* they are directly comparable to each other. As stated earlier, this triplet of normalized tissue fractions constitutes the 3 channels that subsequently guided our multi-subject multi-channel registration algorithm; the outcome of which is presented in the bottom box of Fig. 4.10. Due to the final interpolation (required for deformation), the registered volumes appear slightly softer around the edges. The registration process has brought them all in very good correspondence, save for some normal healthy variation (found mostly at the cortex). The matching process is not biased by a choice of template (or even any artifial way of calculating an intermediary template): all volumes are deformed towards an

average shape that is established by minimizing the viscoelastic deformation energy for each one of them.

4.3.6 Average MSMT and q-space signal template

According to the deformation fields (and their local Jacobian matrices for each voxel) as obtained from the registration, the MSMT representations of each subject were deformed and subjected to PPVF retransformation in each voxel. We provide a whole-brain overview of axial MSMT maps for the template in Fig. 4.11 and Fig. 4.12. The average MSMT template's normalized relative tissue fractions were simply obtained by averaging the deformed normalized

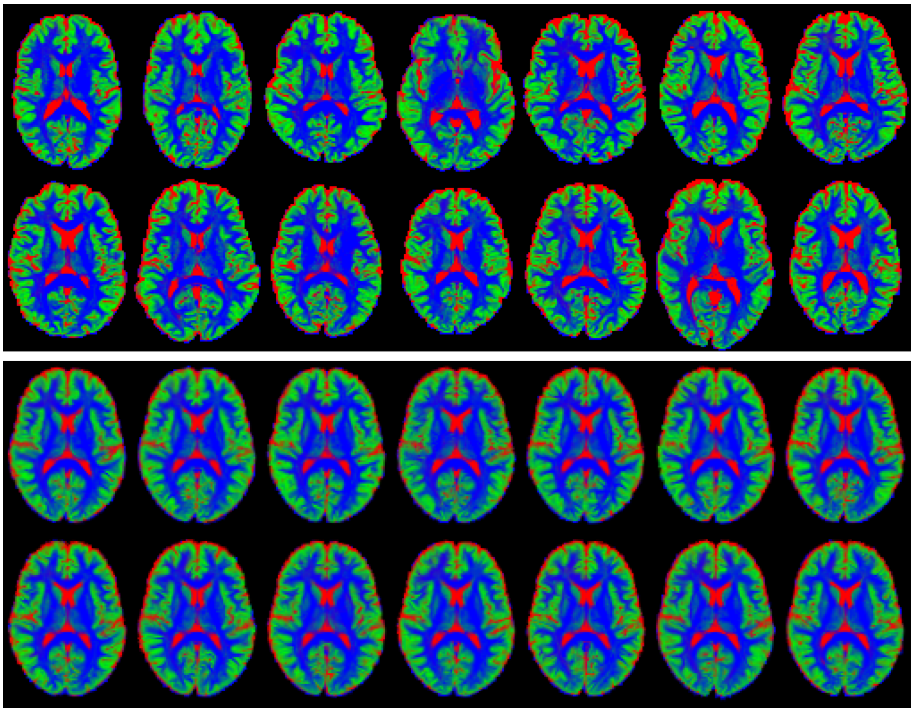

Figure 4.10: Results of the multi-subject multi-channel registration of the normalized tissue fractions. *Top box:* the 14 subjects before registration. *Bottom box:* the 14 deformed subjects after registration. All presented axial slices are located in the same fixed plane through the global space. Note that, due to the 3-dimensional nature of the registration process, anatomical structures are able to move in and out of this visualized plane.

fraction volumes. For the construction of a DEC WM map of the template, a specific issue has to be taken into account: the original set of $D = 300$ directions is transformed by a different Jacobian matrix in each voxel of each subject; *i.e.* the angular WM fractions for a single voxel over the different subjects don't allow for direct averaging. One way of dealing with this could be to evaluate the MSMST representation of the deformed (and PPVF retransformed) volumes for a (large) set of gradient directions, and refitting an MSMT represention, *e.g.* again for the set of $D = 300$ directions. This is requires an extra computational investment, is less accurate (it boils down to some sort of regridding) and is simply unnecessary; as we already have an elaborate separation into tissue fractions. A better solution is to calculate the DEC directly from the complete set of 14×300 fractions and their corresponding directions' individual DECs. This "overall" DEC is then normalized as before and weighted by the (average) normalized WM fraction, as defined in Eq. (4.18). The DEC WM maps as shown in Fig. 4.11 and Fig. 4.12, are obtained by this latter strategy.

Finally, in Fig. 4.13, we present a view on the actual average q-space signal template as resulting from the MSMT representation of the 14 combined subjects. The q-space signal visualizations are overlaid on the template's DEC WM for spatial reference. As the b-value increases, so does the angular contrast. Especially at the higher b-values, one can still clearly spot the interdigitating crossing between the corticospinal tract and the corpus callosum (the DEC WM also provides helpful guidance), as well as partial voluming with the nearby superior longitudinal fasciculus, and also the cingulum bundle. These results, combined with an overall anatomical plausibility of the MSMT maps in Fig. 4.11 and Fig. 4.12, indicate that the registration algorithm guided by the normalized MSMT tissue fractions and combined with a subsequent PPVF retransformation, has performed quite well at matching the different subjects' anatomies in the complete 6-dimensional spatio-angular multi-shell domain.

4.4 Discussion

As the introduced methods' properties and the results are already mostly discussed in the previous sections, we will in this section provide some additional brief brainstorming on future opportunities for research.

4.4.1 Quantitative MSMT

The fractions resulting from the MSMT fit can readily serve quantitative purposes. For group studies or comparisons across different subjects in general,

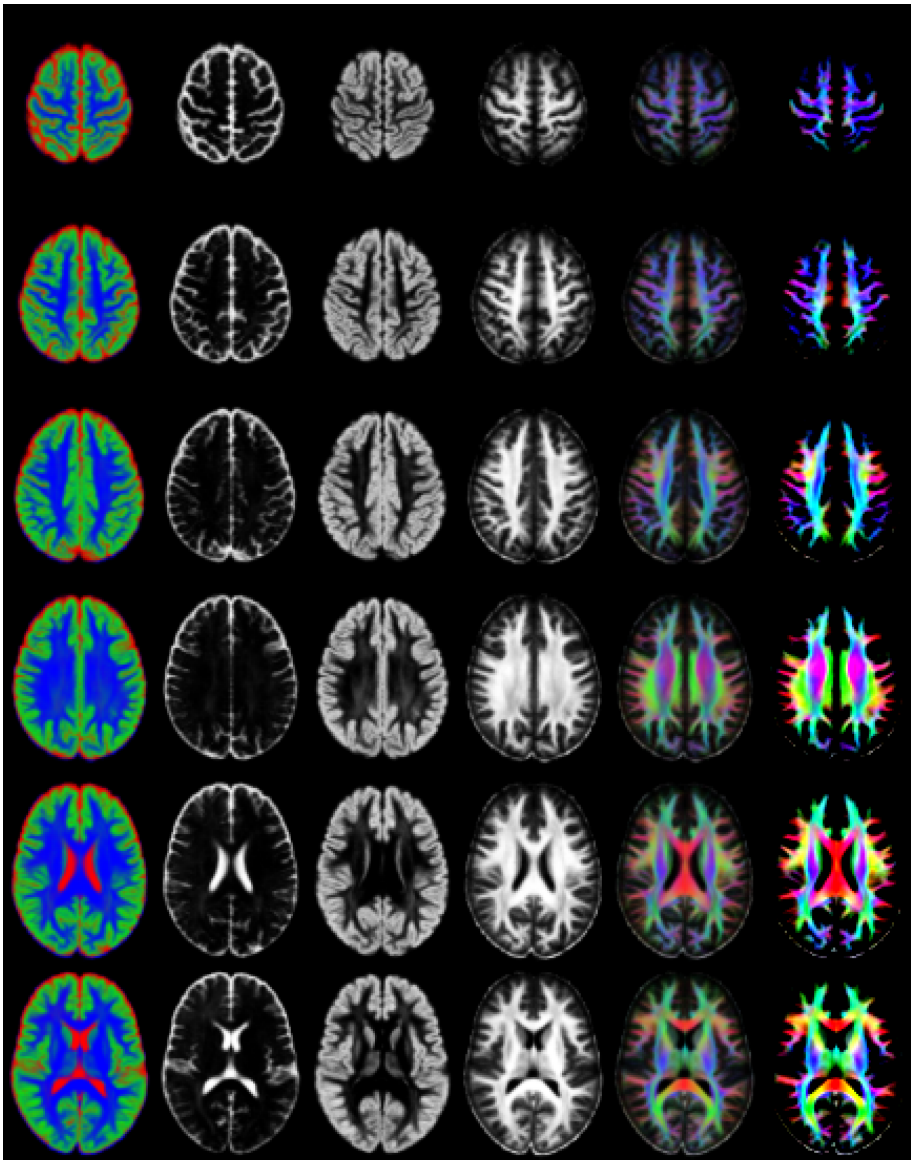

Figure 4.11: Maps obtained from the MSMT representation, visualized at every fourth axial slice for the top half of the average template. From left to right: normalized CSF-GM-WM fractions encoded by RGB; normalized CSF fraction; normalized GM fraction; normalized WM fraction; DEC WM; high contrast DEC WM. The latter one is simply the DEC WM shown at a higher contrast, for the sake of enhanced clarity.

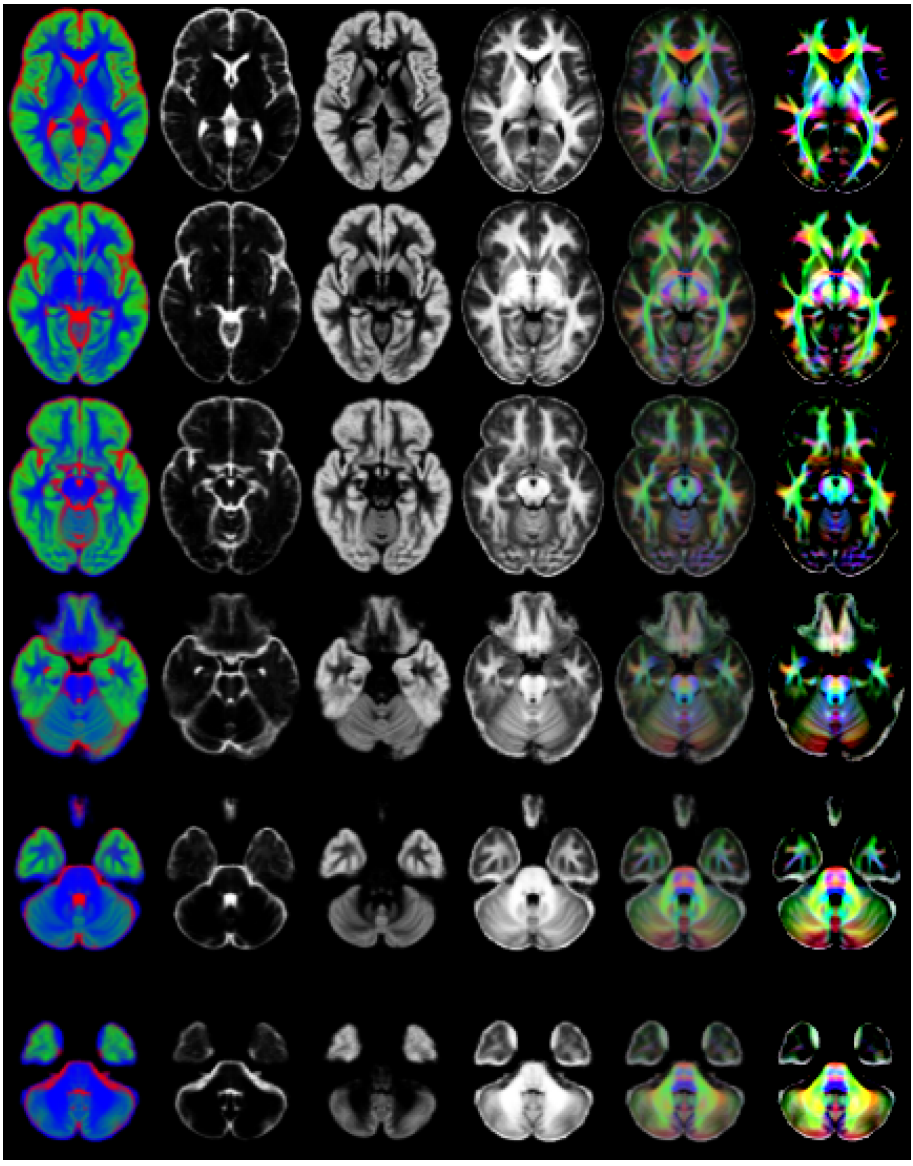

Figure 4.12: Maps obtained from the MSMT representation, visualized at every fourth axial slice for the bottom half of the average template. From left to right: normalized CSF-GM-WM fractions encoded by RGB; normalized CSF fraction; normalized GM fraction; normalized WM fraction; DEC WM; high contrast DEC WM. The latter one is simply the DEC WM shown at a higher contrast, for the sake of enhanced clarity.

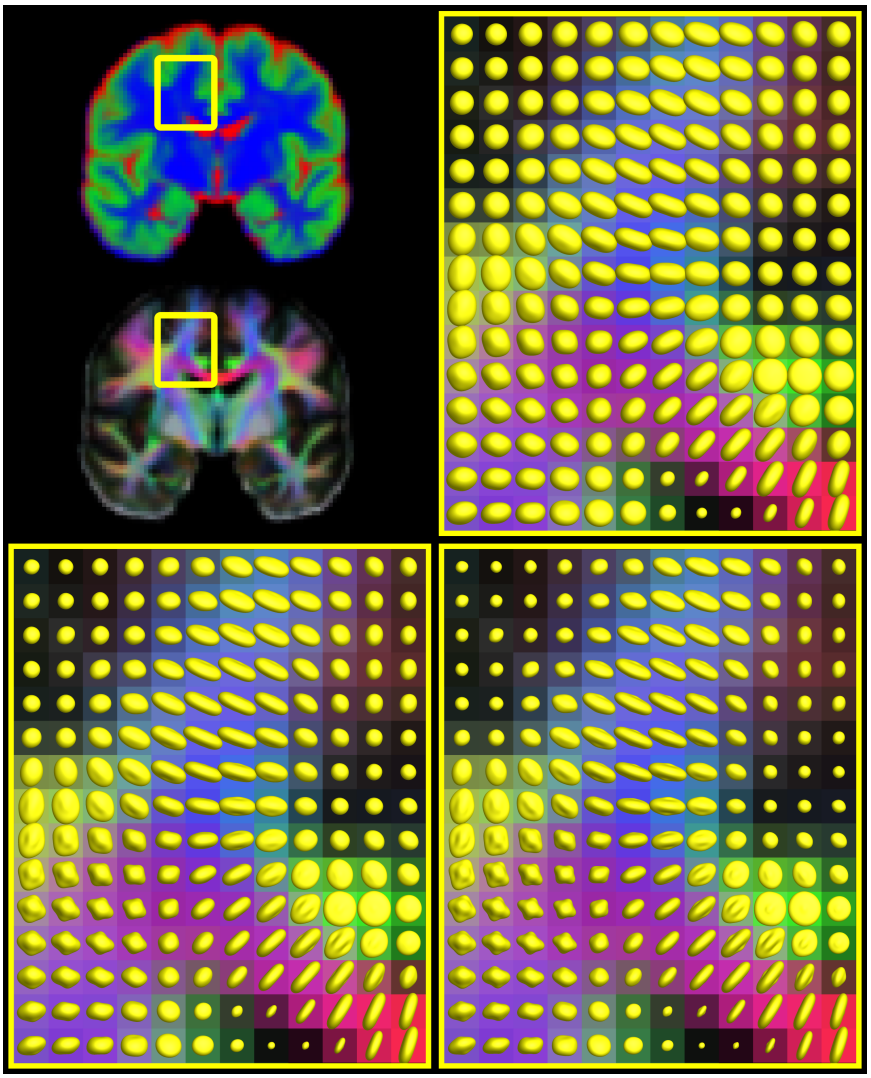

Figure 4.13: The q-space signals of the average template. *Top left:* coronal slice showing the normalized tissue fractions and the DEC WM map, both indicating the location of the zoomed region. *Top right:* $b = 1000\,\mathrm{s/mm^2}$ shell. *Bottom left:* $b = 2000\,\mathrm{s/mm^2}$ shell. *Bottom right:* $b = 3000\,\mathrm{s/mm^2}$ shell. Note that, although the backdrop provides the DEC WM map for spatial reference, the visualized signal results from all three tissue classes. Furthermore, the signals were normalized by the template's $b = 0$ volume. Finally, the three visualizations are each scaled by a (spatially constant) factor for optimal use of space; the actual magnitudes relate to each other in a similar way as they do in Fig. 4.3.

it should be taken in mind that such a quantitative comparison is only valid if a *single fixed* set of tissue response functions is employed for all subjects (as was done in this work). As noted before, the MSMT fractions are transformationally invariant. At the core, this results from the fact that the MSMT representation splits up the geometry of tissue from the actual fractions. Hence, the PPVF retransformation is able to manipulate only the geometry, yet leaves the fractions intact. As such, by applying PPVF retransformation when mapping several DWI datasets by registration, we are actually able to fully *eliminate* all differences related to the geometry. The problem with several classic (DTI) measures, *e.g.* FA, is that they mix up information related to geometry with other properties (Zhang et al., 2012; Jones et al., 2013). Although these measures might be very sensitive, they lack massively in specificity. After registration and PPVF retransformation, however, *differences resulting from direct voxelwise comparisons* can no longer be attributed to a difference in geometry. As such, an important gain in specificity is obtained for differences in properties such as FA (Basser, 1995) and other "shape" measures (Westin et al., 1997). However, if the data allows a good MSMT fit, the fractions themselves should provide a much more direct way of assessing transformationally (*i.e.* geometry-)invariant properties. Our MSMT representation also provides an insight into which measures from other models could qualify as transformational invariants (*e.g.* the ND from NODDI). The MSMT representation itself can even be used to validate such claims directly on real *in vivo* data, by calculating these measures from MSMT evaluated DWI signals before and after PPVF reorientation (as shown for instance in Fig. 4.9).

Specifically when pathological tissue might be present, it could be worthwhile to investigate the residuals of an MSMT fit that employs response functions from the healthy tissue classes of the subject. These residuals could possibly directly highlight areas that cannot be explained "in terms of" healthy tissue of the subject itself. Furthermore, the MSMT framework is of course not limited to the response functions that were used throughout this particular work. If one is able to obtain/design response functions specific to certain pathological tissue classes or effects, their fractions from an MSMT fit could prove to be a very specific, quantitative biomarker. One has to bear in mind, though, that the requirements on the acquisition (*e.g.* the number of required shells) will typically be higher if *extra* tissue classes are added to the MSMT framework.

As mentioned, the response functions do not necessarily have to be measured from the data itself. Existing models of tissue microstructure can be used (or new ones designed) to represent certain compartments. We envision that, similarly to NODDI, a model based CSF IVF could be combined with an intracellular AVF (based on a stick model) and an extracellular AVF (based on a tensor model). In comparison to the original NODDI model, this might

provide for a more "free-form" encoding of the angular dispersion pattern. The feasibility and acquisition requirements for such hypothetical MSMT variants remain to be tested though. The estimation procedure might also need adaption to suit such new settings. Finally, it is worth mentioning that such model-based responses should not assume a mono-exponential DWI signal decay (unless for an isotropic CSF compartment), as this has been shown to cause serious fitting problems for real multi-shell data (Jbabdi et al., 2012).

4.4.2 Measuring response functions

When measuring the response functions directly from the data, as we did in the current work, the challenge lies in finding a decent number of voxels that can quite reasonably be expected to contain "pure" examples of the respective tissue classes. We employed 3 very *ad hoc* pipelines, which resulted in decent outcomes. For the set of common tissue classes (CSF, GM, WM), and data at current typical spatial resolutions on the order of $2 \times 2 \times 2\,\mathrm{mm}^3$, the GM response function is undoubtedly the most difficult to obtain. We showed that an *ad hoc* pipeline can still find a good amount of voxels; there was even a margin to be more restrictive.

A clever trick to obtain an accurate single fiber WM response function in the context of CSD was recently proposed by Tax et al. (2014). The overall idea is to start with a very rough estimate, use this in the fitting (CSD) step, and employ the result itself to obtain a better selection of voxels to measure the response function from (in this case voxels with a single significant peak in the FOD). The procedure is iterated until convergence. This strategy can easily be extended for compatibility with our MSMT representation. Rough estimates of the response functions could be obtained from generous *ad hoc* pipelines. Using these response functions for an MSMT fit should already provide a reasonable view on the (relative) magnitudes of the fractions. For each tissue class, a certain amount of voxels with very high (relative) fractions could be selected (for AVFs, an extra selection to obtain voxels with a single population should be added, akin to the strategy of Tax et al. (2014)). These voxels then provide the data for recalibration of the response functions in the next iteration. Such a strategy could render the complete MSMT fitting process much more automated, user-independent and bias-free (at an additional computational cost).

4.4.3 Fitting strategies

While the fitting procedure presented in this work proved to be quite robust for obtaining the desired *sparse presence of tissue populations* (and not a sparse

appearance *per se*), other fitting strategies could of course be designed and tested. The major advantage of our strategy, is that it obtains a result almost fully complying to certain constraints, yet still in a very fast manner that only requires a few iterations of a simple least squares fit. However, some straightforward simple modifications/extensions can be thought of.

While raising the value of regularization weight λ resulted in a virtual constraint (without any further unwanted effects), we could in a final iteration (after the convergence of the original fitting procedure) even truly enforce this constraint. Due to our discrete representation, this proves to be very easy: just remove the fractions that were, after convergence, assumed to be zero from the system and perform a final fit only for the remaining (non-zero) fractions. In such a final fit, the regularization term associated to λ is simply absent. Furthermore, κ might at this point even be lowered a bit, since a robust estimation of the non-zero (region of) fractions is already established and enforced upon this final iteration.

Another small improvement is possible for the definition of the Laplacian matrix. The current definition (Eq. (4.16)) assumes a perfectly regular grid. While most directions in a large set obtained from electrostatic repulsion will feature 6 neighbors, it is unavoidable that some will only have 5 and quite reasonable that some have 7 (Saff and Kuijlaars, 1997). While this fact by itself is correctly taken into account in Eq. (4.16), the indirect consequence of not all pairwise distances between neighbors being equal is not. For a large set of directions, we expect this issue not to affect the fitting outcome in any significant manner though; but the Laplacian can always be calculated slightly more accurately by making a few small adjustments, as detailed in (Peter et al., 2007).

Apart from such minor tweaks, bigger adjustments (and improvements) might be possible. We envision that the mechanisms described in Tournier et al. (2013b) can be adjusted for our MSMT representation's fitting. This could lower the acquisition requirements in terms of angular resolution, and might deal more effectively with the presence of Rician noise. For single-shell data, one can of course still resort to Yap and Shen (2013) to separate the WM fraction from an overal isotropic fraction (that intends to capture both CSF and GM). We should note, however, that single-shell data combined with one or more $b = 0$ measurements should already provide enough information for a data-driven separation of at least one isotropic fraction (*e.g.* CSF) from the WM fraction, without requiring the artificial assumption of sparse appearance of (angular WM) fractions. In general, we look forward to advances that could robustify or otherwise improve the fitting procedure of MSMT representations.

4.4.4 PPVF out of the box

Finally, we would like to open up the discussion on PPVF retransformation once again. At first sight, it might seem that PPVF retransformation allows the ultimate freedom in the angular domain required to cope with any imaginable spatial transformations. On the other hand, we still assume that all populations within a voxel are bound to stay (spatially) together throughout the transformation. Imagine two interdigitating bundles: if one bundle were to be translated relative to the other, this would currently spatially "drag" the crossing region along (as if both bundles are glued to each other in the crossing zone), potentially impacting the other bundle that should stay in place. Shouldn't the directional WM fractions related to the moving bundle be able to spatially "detach" from the ones related to the fixed bundle; *i.e.* shouldn't they be able to migrate to other voxels? The current logic of keeping the contents of a voxel together as a whole arises mostly from the classical paradigm of a single 3-dimensional spatial displacement field modelling the complete transformation. However, it doesn't seem unreasonable to allow more freedom. This would, however, require a completely different paradigm for dealing with the problem of DWI registration. Since our MSMT representation already splits up the contents of each voxel in the necessary logical building blocks, it might serve as a good basis for (reasoning on) such potentially game-changing developments.

4.5 Conclusion

The discrete MSMT representation and the PPVF retransformation strategy provide an interesting tandem which allows for an intuitive representation and direct manipulation of DWI data. The general framework is open to a plethora of possible modifications that can serve diverse purposes. We look eagerly forward to future research that comes up with creative and useful modifications or applications.

Acknowledgments

The open access datasets used in this work (14 subjects) were provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

Chapter 5

Variability of TDI

This work consists of what is in essence a very simple experiment. The results were presented on three occasions (respectively two poster presentations and an oral presentation):

- Track-density Imaging + Noise: when Super-resolution Quality does not yield Accuracy. (4th annual meeting of the ISMRM Benelux Chapter; Leuven (Belgium), 2012)
- Track-density Imaging & Noise: when Super-resolution Quality does not yield Accuracy. (20th annual meeting of the ISMRM; Melbourne (Australia), 2012)
- How reliable are findings from Track Density Imaging? (29th annual meeting of the ESMRMB; Lisbon (Portugal), 2012)

In the light of the growing popularity of the TDI technique at the time, my obviously rather provoking choice of titles and corresponding message easily gathered a lot of attention. The work itself was received quite well, and resulted in several awards:

- **Nominated for the Best Poster Award (top 8)** for "Track-density Imaging + Noise: when Super-resolution Quality does not yield Accuracy." (4th annual meeting of the ISMRM Benelux Chapter; Leuven (Belgium), 2012)
- **Nominated for the Diffusion Study Group Traditional Poster Award (top 4)** for "Track-density Imaging & Noise: when Super-

resolution Quality does not yield Accuracy." (20th annual meeting of the ISMRM; Melbourne (Australia), 2012)

- **Summa Cum Laude Merit Award (top 3% of abstracts)** for "Track-density Imaging & Noise: when Super-resolution Quality does not yield Accuracy." (20th annual meeting of the ISMRM; Melbourne (Australia), 2012)
- **Certificate of Merit Award (top 10 of abstracts)** for "How reliable are findings from Track Density Imaging?" (29th annual meeting of the ESMRMB; Lisbon (Portugal), 2012)

The following chapter represents an adaption of the original conference abstracts; it aims to retain their brevity, choice of wording and overal clear message.

5.1 Introduction

Track-density imaging (TDI) (Calamante et al., 2010, 2011) is a recently proposed technique that aims to obtain super-resolution anatomical images from (lower resolution) diffusion weighted imaging (DWI) datasets. A typical TDI volume is basically a discretisation of a high number of tracks (resulting from fiber tracking), by counting the number of tracks in each voxel of a high resolution grid. Meaningful anatomical structures beyond the original voxel-scale might then be recovered. In this work, the effect of varying noise in the original data on the final TDI outcome was investigated *in vivo* and *in silico*, in terms of absolute intensities as well as recovered or discovered patterns/structures.

5.2 Materials and Methods

5.2.1 Data

A single healthy subject was scanned in a Siemens 3T scanner using a twice-refocused spin-echo (TRSE) sequence (Reese et al., 2003) at a $2.5 \times 2.5 \times 2.5\,\mathrm{mm}^3$ isotropic voxel size. In addition to 10 non-DWI volumes (which were averaged), 75 DWI volumes were acquired along uniformly distributed gradient directions and using a b-value of $2800\,\mathrm{s/mm}^2$. For the *in silico* data, the same dataset as specified in Calamante et al. (2011) was employed: "Phantom A", obtained from the *Numerical Fibre Generator* software (Close et al., 2009), sampled at 60 gradient directions for a b-value of $3000\,\mathrm{s/mm}^2$. The voxel size was defined to be $2.5 \times 2.5 \times 2.5\,\mathrm{mm}^3$, in order to be consistent with the *in vivo* data.

5.2.2 Bootstrapping and TDI

There were two main steps to be performed for this particular experiment: bootstrapping of the raw data, and subsequent automated application of the complete TDI pipeline. A schematic overview of the whole procedure is provided in Fig. 5.1.

First, 100 different noise realizations of both the real and phantom data were generated. For the real data, this was achieved using the residual bootstrap with correction for leverage, as detailed in Jeurissen et al. (2011). For the phantom data, 100 realizations of random Rician noise at a signal-to-noise ratio of 17 were simply added to the noise-free simulated signals. Next, TDI was applied to all 100 real and 100 phantom bootstrapped datasets, as a fully automated pipeline.

This pipeline starts by applying constrained spherical deconvolution (CSD) (Tournier et al., 2007), using a spherical harmonic order of 8. Subsequently, 6 million tracks are generated for each dataset by probabilistic whole-brain (or "whole-phantom") fiber tractography. These parameters were used: stepsize = 0.2 mm, minimum radius of curvature = 1 mm, minimum track length = 15 mm (5 mm for the phantom), maximum track length = 300 mm in both directions from the seed point (*i.e.* bidirectional tracking was performed), minimum FOD amplitude to initiate tracks = 0.2. A track is terminated if the encountered FOD amplitude drops below 0.1, or when the track leaves a predefined brain mask (this mask was made in advance and is kept fixed for all runs of the pipeline; *i.e.* the mask does not depend on the varying noise). Finally, the number of tracks in each voxel of a new $0.5 \times 0.5 \times 0.5\,\mathrm{mm}^3$ isotropic grid is counted. This results in high resolution TDI volumes (super-resolution: 2.5 mm $\rightarrow$ 0.5 mm isotropic). The complete pipeline was applied fully automatically to all 100 bootstrapped datasets by employing the available functionality in the *MRtrix* package (Tournier et al., 2012).

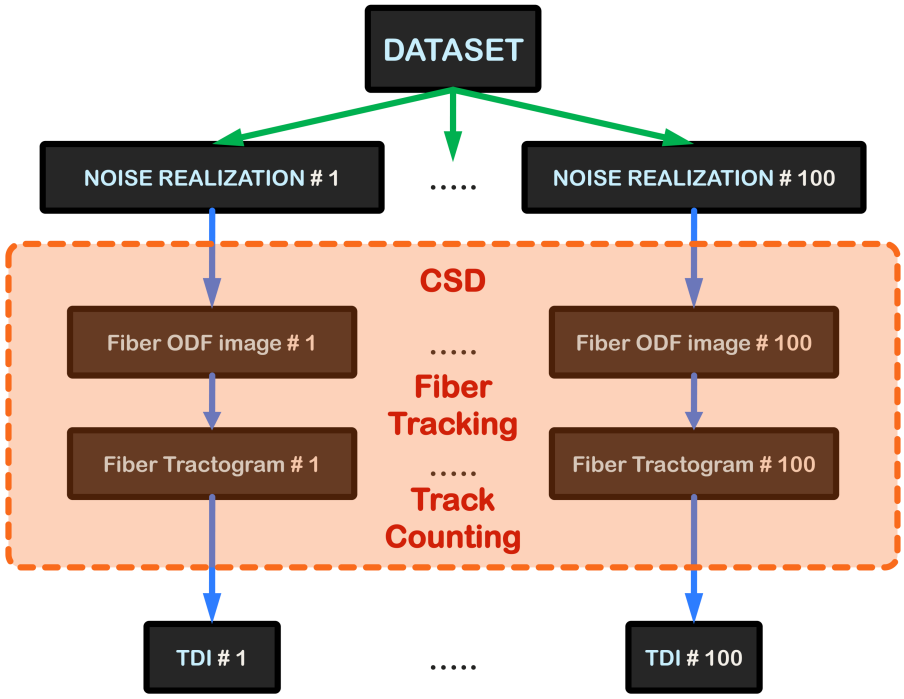

Figure 5.1: A schematic overview of the procedure to obtain 100 TDI volumes that only differ in function of the original random noise.

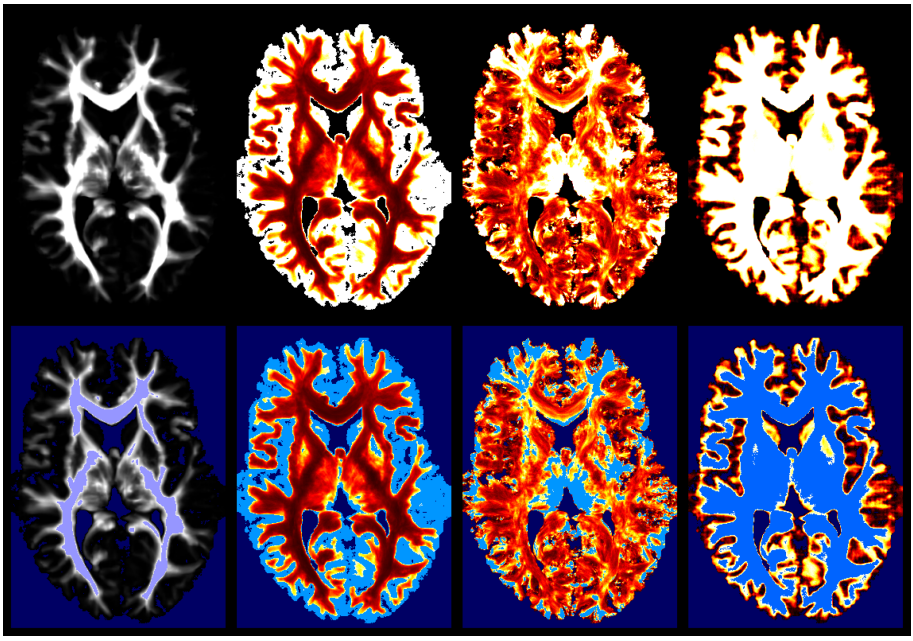

Figure 5.2: *From left to right:* Maps of μTDI (0–300), COV (0–1), COD (0–15) and #R (0–100). The *bottom row* highlights the window caps in *blue*.

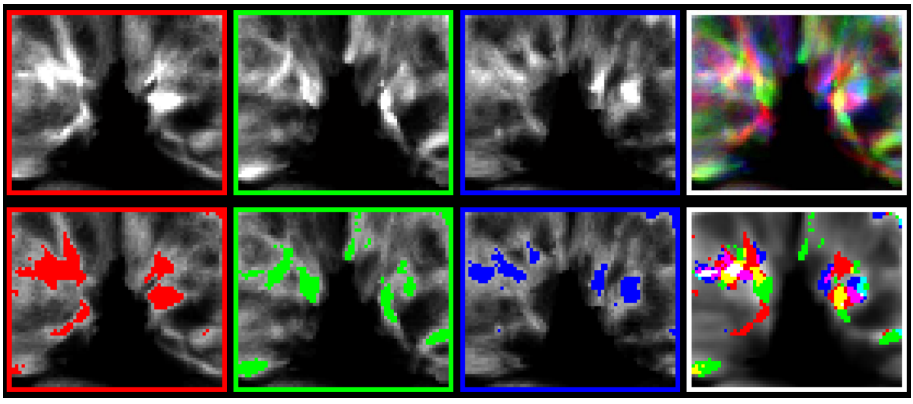

Figure 5.3: *Top row:* the thalamus region for 3 noise realized TDI volumes separately, and combined using RGB color channels. *Bottom row:* ROIs resulting from a fixed absolute threshold, and their combination overlaid on μTDI using RGB color channels.

5.3 Results

Maps of voxel-wise statistics over the 100 *in vivo* TDI volumes are presented in Fig. 5.2. The mean TDI (μTDI) is obtained by calculating the voxel-wise average across the 100 TDI volumes. The coefficient of variation (COV) is defined as σ/μ (where σ is the standard deviation, and μ is the average). The coefficient of dispersion (COD) equals σ^2/μ. Finally, we also defined the "reachability count" (#R) for each voxel as being the number of TDI volumes (out of 100) where the TDI intensity is strictly greater than 0 (*i.e.* the number of massive tractograms that actually reached a particular voxel with at least a single track). All maps in Fig. 5.2 are windowed for optimal visual contrast; only for #R, the full 0–100 range is shown. For the latter, the window caps indicate voxels that were reached in none/all TDI volumes. The COV is the highest at the cortex

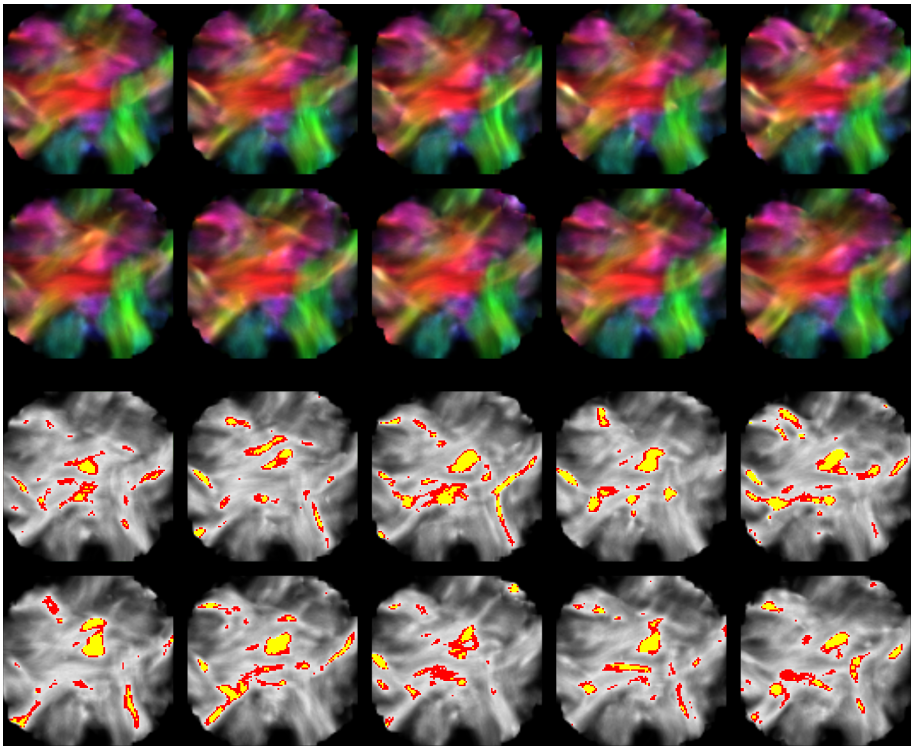

Figure 5.4: *Top rows:* DEC TDI maps of a slice, for 10 of the *in silico* bootstraps. *Bottom rows:* two fixed absolute thresholds on TDI intensity, overlaid on the TDI maps corresponding to the *top rows'* DEC TDI maps.

and in subcortical deep gray matter structures such as the caudate nucleus, the putamen and the thalamus. This corresponds mostly to regions of lower μTDI, and #R below the full 100. The COD map shows quite a distinct pattern. The thalamus (amongst other regions) shows up very brightly. The COD was larger than 1 in *all voxels*. A large variation of TDI patterns across different regions of all volumes was observed. An example of such qualitative pattern variations, and their impact on a region-of-interest (ROI) from an absolute threshold, is shown in Fig. 5.3 for the thalamus region. Finally, directionally-encoded color (DEC) TDI maps of a slice are provided for 10 of the *in silico* bootstraps in Fig. 5.4. Two fixed absolute thresholds on the TDI intensity are also visualized; indicating again a significant qualitative variation of patterns (even though the phantom itself consists of quite well-behaved and regular structures).

5.4 Discussion and Conclusion

While the intensity variations due to the noise constitute by themselves already an important concern, the true (potential) pitfall lies in the fact that ***each individual TDI map looks perfectly plausible***. Even though many of the "discovered" patterns might just as well (mostly) be caused by noise (rather than true anatomical structure), ***they do not give the traditional impression of noise*** (*i.e.* a random high-frequency pattern); as they are a form of *structured* noise resulting from the complex interactions between the measurement noise and the pipeline leading to the final TDI image! Resulting "blobs" in the thalamus (Fig. 5.3) can thus easily be mistaken for true, local anatomical features. Noise in the original data has local, as well as "long distance" effects on some of the intensities and patterns that arise. TDI maps are not robust for quantification (as also stated in Calamante et al. (2012d)). The findings of our current work indicate, however, that great caution is also advised when using TDI maps as super-resolution anatomical maps (qualitatively). New strategies (*e.g.* Pannek et al. (2011a)) show promising results in being more robust (but also show less detailed super-resolution *contrast*). The work in Calamante et al. (2012d) effectively opens up the path to an endless family of new track-weighted maps, some of which *might* be more robust; but again this requires further investigation. One of the properties of such track-weighted maps in general (and TDI in particular) is that they easily "hide" the traditional pattern of noise from the user, giving them a potential impression of high *quality* images. *Accuracy*, however, does not come for free: some observed structures might in fact be caused to a greater extent by this "hidden noise", rather than by true anatomy. Tools such as bootstrapping, combined with maps of COV, COD and #R can be of great value as indicators of regions where caution is advised.

Chapter 6

TODI and TOD-based tractography

The initial results of work on this topic were presented (orally) on two occasions:

- Using Track Orientation Distributions to Robustify Probabilistic Tractography. (5th annual meeting of the ISMRM Benelux Chapter; Rotterdam (Netherlands), 2013)
- Robustifying Probabilistic Tractography by using Track Orientation Distributions. (21st annual meeting of the ISMRM; Salt Lake City (United States), 2013)

For the latter contribution, an award was also received:

- **Summa Cum Laude Merit Award (top 3% of abstracts)** for "Robustifying Probabilistic Tractography by using Track Orientation Distributions." (21st annual meeting of the ISMRM; Salt Lake City (United States), 2013)

Further results of a quantitative validation were presented (as a poster):

- Quantitative Validation of TOD-based Tractography by a Tractometer Approach. (ISMRM Workshop on Diffusion as a Probe of Neural Tissue Microstructure; Podstrana (Croatia), 2013)

Finally, the complete work was published in NeuroImage:

- **Dhollander, T.**, Emsell, L., Van Hecke, W., Maes, F., Sunaert, S., Suetens, P.: Track Orientation Density Imaging (TODI) and Track Orientation Distribution (TOD) based tractography. NeuroImage (in press) (2014)

The following chapter is adapted from the latter published article.

Abstract

Ever since the introduction of the concept of fiber tractography, methods to generate better and more plausible tractograms have become available. Many modern methods can handle complex fiber architecture and take on a probabilistic approach to account for different sources of uncertainty. The resulting tractogram from any such method typically represents a finite random sample from a complex distribution of possible tracks. Generating a higher amount of tracks allows for a more accurate depiction of the underlying distribution. The recently proposed method of track-density imaging (TDI) allows to capture the spatial distribution of a tractogram. In this work, we propose an extension of TDI towards the 5D spatio-angular domain, which we name track orientation density imaging (TODI). The proposed method aims to capture the full track orientation distribution (TOD). Just as the TDI map, the TOD is amenable to spatial super-resolution (or even sub-resolution), but in addition also to angular super-resolution. Through experiments on *in vivo* human subject data, an *in silico* numerical phantom and a challenging tractography phantom, we found that the TOD presents an increased amount of regional spatio-angular consistency, as compared to the fiber orientation distribution (FOD) from constrained spherical deconvolution (CSD). Furthermore, we explain how the amplitude of the TOD of a short-tracks distribution (*i.e.* where the track length is limited) can be interpreted as a measure of track-like local support (TLS). This in turn motivated us to explore the idea of TOD-based fiber tractography. In such a setting, the short-tracks TOD is able to guide a track along directions that are more likely to correspond to continuous structure over a longer distance. This powerful concept is shown to greatly robustify targeted as well as whole-brain tractography. We conclude that the TOD is a versatile tool that can be cast in many different roles and scenarios in the expanding domain of fiber tractography based methods and their applications.

6.1 Introduction

Since the advent of diffusion weighted imaging (DWI) in the mid-1980s, we have come a long way towards mapping the structural network of the (human) brain *in vivo* and non-invasively (Jones, 2010), leading to better insight in its complexity and eventually giving rise to a whole new field of connectomics (Hagmann, 2005; Sporns et al., 2005; Hagmann et al., 2010). A key development in achieving this has certainly been the introduction of the concept of fiber tractography at the end of the previous millenium (Basser, 1998; Basser et al., 2000; Jones

et al., 1999b; Mori et al., 1999; Poupon et al., 1999; Lori et al., 1999; Conturo et al., 1999). Throughout the past decade, tractography has proven to be a powerful tool that comes with many advantages as well as important limitations (Jbabdi and Johansen-Berg, 2011). Current state-of-the-art algorithms, that can handle complex fiber architecture (*i.e.* so-called "crossing fibers") as well as take on a probabilistic approach to account for different sources of uncertainty, are now widespread and available for public use, *e.g.* Tournier et al. (2012); and were proven to be far superior to fundamentally limited techniques based on diffusion tensor imaging (DTI) for critical applications such as neurosurgical planning and navigation (Farquharson et al., 2013). Given the orientational information in each voxel (*e.g.* a fiber orientation distribution (FOD) from constrained spherical deconvolution (CSD), Tournier et al. (2007, 2008)) and several constraining parameters (*e.g.* minimum and/or maximum track length, limited curvature, minimum FOD amplitude threshold, a mask, a seed region, ...), a tractography algorithm yields a tractogram: a finite number of accepted tracks; a random sample from a distribution of possible tracks. The more tracks are generated, the more representative the tractogram becomes regarding this distribution.

Recently, track-density imaging (TDI) (Calamante et al., 2010, 2011) was proposed as a means to obtain super-resolution from DWI data by exploiting the continuous nature of dense tractograms consisting of a very large number (*e.g.* millions) of tracks. It does so by employing a strategy very similar to that of a common histogram: in each voxel of a (potentially high resolution) 3D grid, the number of tracks is simply counted. As the amount of tracks that can be generated is virtually unlimited, the underlying distribution can be approximated up to any detail, justifying the use of super-resolution. As such, TDI is an excellent tool to obtain a "classical" 3D image space representation of the spatial probability density function (PDF) of complex track distributions. In an attempt to capture and visualize some additional angular information, directionally-encoded color (DEC) TDI assigns to each track passing through a certain voxel a color based on the local orientation of the track (conforming to the color scheme widely used for DEC fractional anisotropy (FA) maps in DTI, *e.g.* red: mediolateral, green: anteroposterior, and blue: superoinferior). These colors (RGB values) are then summed in each voxel, resulting in the DEC TDI map. While the global pattern of colors does provide an additional visual cue for easier localization of known structures, it can on a local (voxel) level not distinguish between multiple fiber populations, nor describe other complex structure. This particular color scheme even lacks specificity in describing a *single* direction (*i.e.* even in DEC FA maps from DTI): in general, an RGB triplet can correspond to four different directions (when all R-, G- and B-values are nonzero). Therefore, the additional DEC in the context of TDI has no major use beyond being a practical visualization tool.

In this work, we introduce track orientation density imaging (TODI), a technique which aims to reconstruct a complete description of the track orientation distribution (TOD) in each voxel. We will describe the formalisms of TODI as a generalization of TDI to the entire 5D spatio-angular domain. To maintain full compatibility with the existing definition of TDI, the same mechanism is employed to map the tracks' spatial distribution: each track delivers a *unit* contribution to each voxel it intersects and thus the final weight of a voxel in the spatial distribution is proportional to its track count. Whereas this choice of using voxels as bins for a histogram-like approach in the 3D spatial domain is a straightforward one, it is less trivial to define such discrete bins in the 2D angular domain, *i.e.* on the sphere. A possible approach to discrete angular binning consists of generating a large set of uniformly distributed directions, *e.g.* by geodesation of an icosahedron or electrostatic repulsion (Jones et al., 1999a). A track could then locally contribute to the direction within the set that best approximates the track's (tangent) direction (Pannek et al., 2012). Such a strategy conceptually equates to defining a set of angular bins as resulting from the Voronoi tessellation (on the sphere) of the set of generated directions. Typically, most of these bins would be hexagonally shaped, though some would inevitably be pentagons (Saff and Kuijlaars, 1997), and using electrostatic repulsion might in practice even result in some heptagons. To simply avoid possible bias caused by any particular choice of angular bins, we rather choose to estimate the angular part of the distribution by use of a kernel density estimate (KDE). In TODI, we will represent the TOD in each voxel using a set of spherical harmonics (SH) basis functions and construct it by continuous integration of a spherical point spread function (PSF) along the intersecting part of each track: the resulting contribution of a track to a voxel is a full angular function. As for the PSF, we will first consider a SH delta function and further refine this choice to an apodized delta function (Raffelt et al., 2012a) to avoid the Gibbs truncation artifacts associated with the SH delta function's definition. More information on the construction of such apodized delta functions can be found in Appendix A. Based on the same principles as discussed in Calamante et al. (2010), TODI is also amenable to spatial super-resolution. Furthermore, we will discuss how these principles are inherently extended to the angular domain, effectively allowing for the complementary case of angular super-resolution. Even spatial *sub*-resolution could prove to be an interesting option in order to investigate the angular distribution of a tractogram over larger spatial volumes as a whole. The most extreme case of such spatial sub-resolution is to fully ignore all spatial information and only consider the angular distribution; *i.e.* as if the whole tractogram were encapsulated in a single spatial bin (*e.g.* a single large voxel).

Although TODI can be applied to *any* tractogram resulting from *any* fiber tractography algorithm, the results presented in this work are focused on its

application to so called "short-tracks" tractograms. The short-tracks strategy was proposed in Calamante et al. (2012c) to mitigate the effect of the TDI map of a "regular" whole-brain tractogram featuring higher intensities in longer tracts[1], caused by those tracts containing more seed points. By imposing an upper limit on the track length, the short-tracks strategy distributes track densities more evenly over the brain. In our interest, this avoids dependence of TOD amplitudes on different relative tract lengths. Additionally, this renders the TOD in most voxels more comparable (qualitatively) to the FOD as resulting from CSD (Tournier et al., 2007, 2008), which is also independent of tract length. Even though the short-tracks TOD in each voxel certainly has, by definition and construction, a *completely different meaning* as compared to the FOD, it is of a similar qualitative nature in the sense that it also features sharp lobes along directions associated with white matter pathways. We will, however, explain how its amplitude can be interpreted as a measure of *track-like local support* (TLS). This consequently renders the short-tracks TOD itself a potentially interesting candidate to guide tractography using existing algorithms that were originally designed to perform FOD-based tractography. We reason that, in a tractography setting, the short-tracks TOD should be able to guide a track along directions that are more likely to correspond to continuous structure over a longer distance. We will explore the idea of such TOD-based tractography and compare it directly to FOD-based tractography. We also take this concept to the next level by employing TOD-based tractography itself to generate a new short-tracks tractogram that can be used to construct yet another TOD. The latter TOD can in its own right be used for the same purpose (*e.g.* again for TOD-based tractography). The spatial and angular properties of these TODs and the FOD, and tractography results based on them, will be extensively compared using healthy human subject data of different qualities (*i.e.* high angular resolution diffusion imaging (HARDI) as well as more DTI-like acquisition schemes), an *in silico* numerical phantom (Close et al., 2009) and a challenging tractography phantom (Poupon et al., 2008, 2010; Fillard et al., 2011).

Finally, the concept of TDI and similar techniques has been extended to a generalized framework of track-weighted imaging (TWI) (Calamante et al., 2012d), where the resulting maps can depend on any property of the tracks or other underlying image data (a similar technique was proposed in Pannek et al. (2011b)). In our current work, we generalize TDI to the 5D spatio-angular domain, obtaining TODI. The fact that TDI can independently be extended to both of these generalizations (TWI and TODI) begs the question if their mechanisms can be integrated with each other to obtain an even more general framework. Because we took special care to design TODI in a very

[1]Throughout this chapter, the term "track" denotes a single streamline as obtained from a tractography algorithm, while "tract" refers to a full white matter bundle/structure (*e.g.* the corticospinal tract).

TWI-compatible way, the answer is a clear "yes" and the resulting framework of track orientation weighted imaging (TOWI) is easy to obtain. Even though an experimental investigation of the concept and possibilities of TOWI is far beyond the focus of the current work, we provide the basic formalisms in Appendix B.

6.2 Materials and Methods

6.2.1 Track Orientation Density Imaging (TODI)

To map the distribution of the tracks in the spatial domain, TODI applies the same definition as TDI: the number of tracks passing through a single voxel directly determines the value or weight of that voxel in the final distribution. This is conceptually similar to a histogram, where the presence of datapoints is counted in discrete bins. However, for the angular part of the distribution we resort to the mechanics of a KDE so as to avoid the definition of discrete bins on the unit sphere. We employ a modified SH basis (Descoteaux et al., 2006) in order to directly obtain a continuous representation of the TOD. Such a basis is defined using the standard SH functions Y_ℓ^m by

$$Y_j = \begin{cases} \sqrt{2} \cdot \mathrm{Re}(Y_\ell^m) & \text{if } -\ell \leq m < 0 \\ Y_\ell^0 & \text{if } m = 0 \\ \sqrt{2} \cdot \mathrm{Im}(Y_\ell^m) & \text{if } 0 < m \leq \ell \end{cases} \tag{6.1}$$

where only the even orders ℓ are used and $j = (\ell^2 + \ell + 2)/2 + m$. It has the useful properties of being real-valued, antipodally symmetric and orthonormal with respect to the inner product. If it is limited up to a certain maximum order n, it consists of $T = (n+1)(n+2)/2$ basis functions.

The SH delta function is a PSF that can be easily obtained in this basis. A SH delta function $\delta_n^z(u)$ up to order n with its central axis of symmetry and maxima along a direction z is obtained by projecting an antipodally symmetric spherical delta function $\delta^z(u)$ into its SH coefficients e_j^z using

$$\begin{aligned} \delta_n^z(u) &= \sum_{j=1}^{T} e_j^z Y_j(u) \\ \iff e_j^z &= \oint_{\mathbb{S}^2} \delta^z(u) Y_j(u)\,\mathrm{d}u = Y_j(z) \end{aligned} \tag{6.2}$$

where $\mathbb{S}^2$ is the surface of the unit sphere. If directions are expressed as pairs of spherical coordinates by $u = (\theta, \phi)$, then the surface element of $\mathbb{S}^2$ can be

written as $\mathrm{d}u = \sin(\theta)\,\mathrm{d}\theta\,\mathrm{d}\phi$. Given a certain direction z, all coefficients e_j^z can be calculated independently of each other and the maximum order n. Any $\delta_n^z(u)$ itself integrates to 1 over $\mathbb{S}^2$ (this integral depends only on $e_1^z = Y_1(z) = 1/\sqrt{4\pi}$). The function $\delta_n^z(u)$ is used as a kernel to indicate the presence of a line element of a track along direction z. Such a line element and the corresponding $\delta_n^z(u)$ for n varying from 2 to 16 are shown in Fig. 6.1 (*first row*).

By linearly parameterizing distance along a track k by ε, a kernel $\delta_n^{z_\varepsilon}(u)$ is assigned to each position ε, with z_ε being the local orientation of the track in ε. To calculate the full contribution $C_n^k(r,u)$ of a single track k to a voxel r, the kernel is averaged along the length that k travels through r. This contribution and its SH coefficients $c_j^k(r)$ can consequently be written as

$$C_n^k(r,u) = \int_{\varepsilon_1}^{\varepsilon_2} \frac{\delta_n^{z_\varepsilon}(u)}{\varepsilon_2 - \varepsilon_1}\,\mathrm{d}\varepsilon \tag{6.3}$$

$$= \sum_{j=1}^{T} c_j^k(r) Y_j(u)$$

$$\iff c_j^k(r) = \int_{\varepsilon_1}^{\varepsilon_2} \frac{Y_j(z_\varepsilon)}{\varepsilon_2 - \varepsilon_1}\,\mathrm{d}\varepsilon \tag{6.4}$$

where ε_1 and ε_2 are the endpoints of the intersecting part of k with r (these are typically the points where the track enters/exits the voxel, but can also be one of the absolute endpoints of the track). An example of a single track intersecting a voxel and the corresponding $C_n^k(r,u)$ for n varying from 2 to 16 are shown in Fig. 6.1 (*second row*). The contribution is a continuous function encoding the full angular footprint of the track, *i.e.* even its curvature, in the voxel. In TDI, the contribution of a single track to a voxel is simply a scalar 1. In TODI, each track also delivers a *unit* contribution to each voxel: the contribution integrates to 1 over $\mathbb{S}^2$, because the kernel is averaged and also integrates to 1 over $\mathbb{S}^2$. In practice, the contribution can be calculated by a summation over track segments of equal length, divided by the number of segments. However, to obtain an accurate representation of the track, the step size (*i.e.* segment length) should be reasonably smaller than the voxel size, or at least small enough to represent the continuous path of the track through the angular domain. If this is not sufficiently the case, smaller segments can be calculated by smooth (*e.g.* Hermite) interpolation.

The final track orientation distribution $\mathrm{TOD}_n(r,u)$ for a voxel r is obtained by a simple summation of the contributions of all tracks to r. The TOD and its

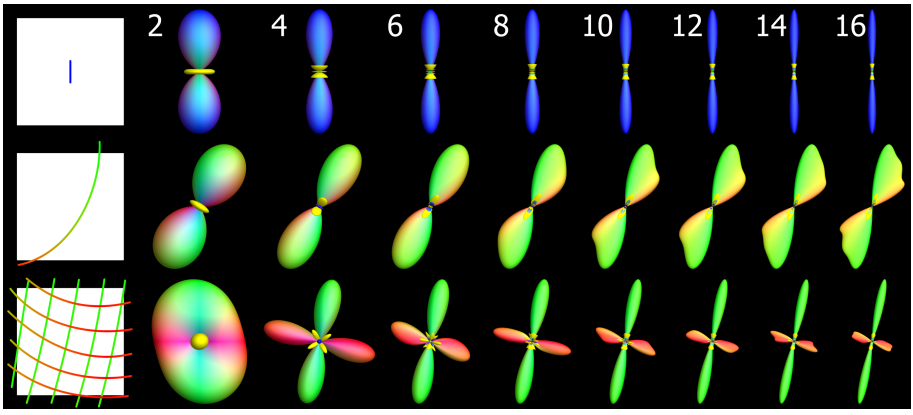

Figure 6.1: TODI using SH delta functions. *First row:* an infinitesimal line element, represented by a SH delta function of increasing maximum SH order. The solid yellow side lobes represent negative values. *Second row:* a part of a track intersecting a voxel, represented by integrating the SH delta function along its length. *Third row:* a group of tracks, represented by summing its tracks' contributions. Negative values and Gibbs truncation artifacts can be seen for all maximum SH orders.

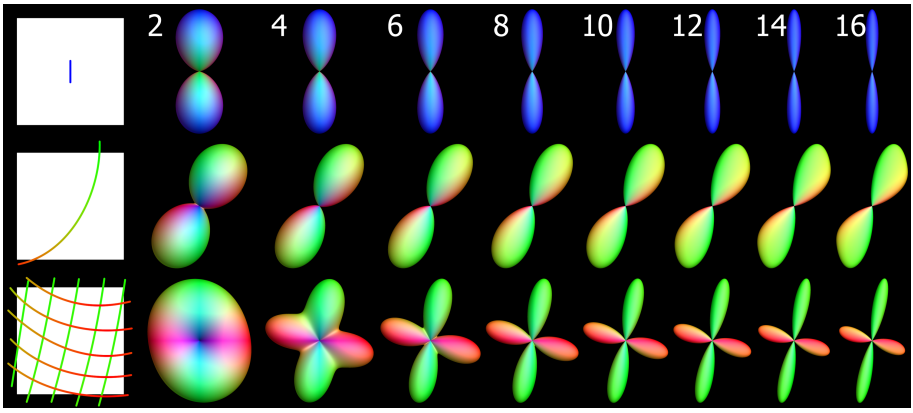

Figure 6.2: TODI using apodized delta functions. *First row:* an infinitesimal line element, represented by an apodized delta function of increasing maximum SH order. *Second row:* a part of a track intersecting a voxel, represented by integrating the apodized delta function along its length. Increased maximum SH order allows for a more accurate description of the track's angular profile. *Third row:* a group of tracks, represented by summing its tracks' contributions. Increased maximum SH order allows for better differentiation of both populations. Note that both the red and green lobes' integrals are *equal* in this example, even though the lobes differ in shape. This integral is however *not* equal to the lobes' volume when the radius is made proportional to the amplitude in such a visualization, making it difficult to appreciate this property.

SH coefficients $t_j(r)$ are thus given by

$$\mathrm{TOD}_n(r,u) = \sum_{i=1}^{K(r)} C_n^{k_i}(r,u) \tag{6.5}$$

$$= \sum_{j=1}^{T} t_j(r) Y_j(u)$$

$$\iff t_j(r) = \sum_{i=1}^{K(r)} c_j^{k_i}(r) \tag{6.6}$$

where $K(r)$ equals the total number of tracks passing through r. An example of 10 tracks intersecting a voxel and the corresponding $\mathrm{TOD}_n(r,u)$ for n varying from 2 to 16 are shown in Fig. 6.1 (*third row*). The TOD in a voxel is a continuous function representing the complete angular distribution of all tracks within the voxel. Because all SH coefficients $t_j(r)$ are defined independently of each other and the maximum order n, the coefficient vector of a TOD_n also contains the coefficients of all TODs of lower maximum order. Integrating a TOD over $\mathbb{S}^2$ results in the TDI value, which simply equals $K(r)$. This information is fully encoded by the $t_1(r)$ coefficient image up to a constant factor: the TDI map equals $t_1(r) \cdot \sqrt{4\pi}$. Put differently, TDI is the special case "TOD_0" of TODI.

6.2.2 TODI with apodized delta functions

Even though the SH delta function is easy to compute, it's not the most optimal choice for the purpose of constructing an angular KDE: the projection of a spherical delta function into its SH coefficients using Eq. (6.2) results in Gibbs truncation artifacts; *i.e.* several positive and negative side lobes, as can be clearly observed in Fig. 6.1 (*first row*). Not only does the accumulation of negative side lobes yield undesired negative TOD values, the complete side lobe patterns also interfere with each other and distort the shape of the contributions and the final angular distribution (Fig. 6.1, *second and third rows*). A better suited choice of kernel is the apodized delta function (Raffelt et al., 2012a) which can, however, not be calculated in an equally easy manner as a SH delta function. More details on the construction of apodized delta functions can be found in Appendix A.

In Fig. 6.2 (*first row*) the apodized delta functions $\widetilde{\delta}_n^z(u)$ for n varying from 2 to 16 are shown, now representing a line element without side lobes. This kernel can be incorporated in the definition of a contribution $\widetilde{C}_n^k(r,u)$ of a track

k to a voxel r, similar to Eq. (6.3). This particular contribution can then be further simplified as follows:

$$\begin{aligned}
\widetilde{C}_n^k(r,u) &= \int_{\varepsilon_1}^{\varepsilon_2} \frac{\widetilde{\delta}_n^{z_\varepsilon}(u)}{\varepsilon_2-\varepsilon_1}\,\mathrm{d}\varepsilon \\
&= \int_{\varepsilon_1}^{\varepsilon_2} \frac{\widetilde{\delta}_n(\theta) * \delta_n^{z_\varepsilon}(u)}{\varepsilon_2-\varepsilon_1}\,\mathrm{d}\varepsilon \\
&= \widetilde{\delta}_n(\theta) * \int_{\varepsilon_1}^{\varepsilon_2} \frac{\delta_n^{z_\varepsilon}(u)}{\varepsilon_2-\varepsilon_1}\,\mathrm{d}\varepsilon \\
&= \widetilde{\delta}_n(\theta) * C_n^k(r,u) \qquad (6.7)
\end{aligned}$$

leading to a straightforward relation between $\widetilde{C}_n^k(r,u)$ and $C_n^k(r,u)$, where the $*$ operator signifies spherical convolution and $\widetilde{\delta}_n(\theta)$ is the apodized delta function of maximum order n with its central axis of symmetry and maxima along the $\theta = 0$ direction (which consequently only varies in function of θ). This elegant result is only possible because of our choice to use the SH delta functions in the original definition (Eq. (6.3)), combined with a property (Eq. (6.14)) that allows for easy rotation of any axially symmetric kernel (see Appendix A). An example of a single track intersecting a voxel and the corresponding $\widetilde{C}_n^k(r,u)$ for n varying from 2 to 16 are shown in Fig. 6.2 (*second row*).

The definition of the final track orientation distribution $\widetilde{\mathrm{TOD}}_n(r,u)$, similar to Eq. (6.5) but using the apodized delta function up to order n as a kernel, can also be simplified:

$$\begin{aligned}
\widetilde{\mathrm{TOD}}_n(r,u) &= \sum_{i=1}^{K(r)} \widetilde{C}_n^{k_i}(r,u) \\
&= \sum_{i=1}^{K(r)} \widetilde{\delta}_n(\theta) * C_n^{k_i}(r,u) \\
&= \widetilde{\delta}_n(\theta) * \sum_{i=1}^{K(r)} C_n^{k_i}(r,u) \\
&= \widetilde{\delta}_n(\theta) * \mathrm{TOD}_n(r,u) \qquad (6.8)
\end{aligned}$$

leading to a similarly elegant relation between $\widetilde{\mathrm{TOD}}_n(r,u)$ and $\mathrm{TOD}_n(r,u)$: a simple spherical convolution with the apodized delta kernel $\widetilde{\delta}_n(\theta)$. This means

that, given a certain n, we only need to calculate the coefficients of $\widetilde{\delta}_n(\theta)$ once. The single final spherical convolution in Eq. (6.8) can be performed by direct multiplication of the SH coefficients of the original TOD with the rotational harmonics (RH) coefficients of the apodized delta kernel (Healy et al., 1998; Tournier et al., 2004, 2007). Note that this derivation is equally valid starting off from any other axially symmetric kernel represented in the SH basis up to a certain order n. An example of 10 tracks intersecting a voxel and the corresponding $\widetilde{\mathrm{TOD}}_n(r,u)$ for n varying from 2 to 16 are shown in Fig. 6.2 (*third row*). Just as before, integrating a $\widetilde{\mathrm{TOD}}$ over $\mathbb{S}^2$ yields the TDI value $K(r)$ and this information is fully present in the first coefficient (or in the "$\widetilde{\mathrm{TOD}}_0$") up to a constant factor. Do note that, unlike the TOD_n, the coefficient vector of a $\widetilde{\mathrm{TOD}}_n$ cannot simply be truncated to obtain the one for a lower maximum order, because the SH coefficients of an apodized delta function depend on its maximum order n (see Appendix A). In this case, the coefficient vector of the TOD should be truncated to the desired order before convolving with the appropriate apodized delta function of equal order to obtain the final $\widetilde{\mathrm{TOD}}$.

6.2.3 Super-resolution and sub-resolution

Because the mapping approach employed by TODI in the spatial domain is equivalent to TDI, it is also amenable to spatial super-resolution based on the principles put forward in Calamante et al. (2010). At the core, these super-resolution properties are a direct consequence of the model of fiber tractography. Because a tractography algorithm in general enforces some constraints on the continuity of the generated tracks, information can be obtained about partial volume fractions within a voxel. Typical constraints include a maximum amount of curvature and a minimum track length. The combination of both guarantees that a track segment can only be present in a given location if it is *supported* by its neighborhood to some extent: directional information in its immediate surroundings may not deviate too much (maximum curvature) and this property should hold over a certain distance (minimum track length) along the track. Per consequence, the edges of the resulting bundles of tracks always feature a nice amount of continuity and mark clear-cut delineations of known white matter tracts, provided that enough tracks are generated to densely fill space inside the tracts. The extra angular information included in the TOD can additionally help to accurately distinguish between several crossing tracts, where a (DEC) TDI map lacks the information content to do so. Regarding the application of these maps, a word of caution is due though. As has been stated in Willats et al. (2012), the actual values (track counts) of super-resolution TDI maps resulting from current typical probabilistic tractograms are not ideal for quantitative analysis due to large *within*-subject and *between*-subject variability; a finding

that is consequently of equal relevance with regard to the absolute amplitudes of the TOD. Furthermore, we have shown ourselves in Dhollander et al. (2012) that using qualitative features from the TDI map in an attempt to discover new structures *beyond* the original voxel-scale can easily lead to overinterpretation, as such structures might be the direct result of noise in the DWI data rather than true anatomy. However, both of these findings are of course a direct consequence of the properties of the applied tractography pipeline; not of the TDI and TODI mapping approaches as such.

Applying the aforementioned super-resolution principles is also possible in the angular domain: the angular content in the TOD should as well only be present when it is *supported* by the neighborhood, under the same constraints put forward by tractography (*e.g.* maximum curvature and minimum track length). These powerful constraints can further add upon possible super-resolved information that might already have been obtained during the reconstruction of the FOD, *e.g.* by enforcing the non-negativity constraint of CSD (Tournier et al., 2007). While increasing spatial resolution is done by decreasing the size of voxels (bins), increasing angular resolution is achieved by increasing the SH order of TODI, *i.e.* by use of a sharper kernel. An increased angular resolution allows for a more accurate representation of even a single track in each voxel (Fig. 6.2, *second row*), and per consequence also improves the description of a population of tracks (Fig. 6.2, *third row*). In this work, we will always start by reconstructing the FOD up to a SH order 8 by employing CSD (super-resolved, if necessary), followed by probabilistic tractography using *MRtrix* (Tournier et al., 2012). Regardless of the chosen spatial resolution for TODI, we will consistently reconstruct $\widetilde{\mathrm{TOD}}_{16}(r,u)$. While the increase from order 8 to 16 (*i.e.* 45 to 153 parameters) may seem like a big leap, we should stress that the choice of an apodized delta function on the other hand causes some *loss* of angular resolution. In fact, the "sharpness" of order 16 results when using apodized delta functions (Fig. 6.2, constructed using the algorithm and parameters provided in Appendix A) is comparable to using SH delta functions (Fig. 6.1) at a maximum order of 8 to 10; but without the Gibbs artifacts. Hence, the extra parameters only partially result in increased angular resolution, while being mostly invested in accuracy through mitigation of the Gibbs artifacts. For the sake of compact notation, we will be referring to our $\widetilde{\mathrm{TOD}}_{16}(r,u)$ results informally by use of the shorter term "TOD" from this point onwards.

Furthermore, even the concept of spatial *sub*-resolution is a theoretical possibility, as our accurate angular mapping approach still allows to fully describe the increasingly complex angular fiber architecture of larger spatial regions. The spatial bins don't even have to take on the form of (cubic) voxels; they can also result *e.g.* from the delineation of specific structures. Rather than a more detailed description (as resulting from spatial super-resolution), this could

provide a better general overview: TODI can indeed be applied at several spatial scales for increased insight in the architecture of a complex tractogram. In this work, we will only provide an illustrative example of the most extreme case of such spatial sub-resolution: effectively ignoring all spatial information while only capturing the angular content of a tractogram's distribution using the KDE mechanism of TODI. This can be achieved by defining a single large spatial bin that includes the whole tractogram, or in the classical sense where bins take on the shape of voxels, by encapsulating the tractogram in a single big voxel. When such a method is applied to a whole-brain tractogram, the result could provide an insight in its complex global angular architecture.

Finally, angular *sub*-resolution is of course also perfectly possible: the most extreme case is TDI, which simply ignores the angular information and only maps the spatial distribution.

6.2.4 Short tracks and Track-like Local Support (TLS)

A characteristic feature of TDI maps of "regular" whole-brain tractograms is their very high dynamic range. This can be explained by longer tracts receiving more seed points when these seeds are distributed uniformly over the brain, resulting in more tracks and thus an associated higher track density. Another way to understand this is the fact that tracks originating from seed points in these longer tracts are able to run along greater lengths and therefore contribute to more voxels. In order to mitigate the dependence of TDI on tract length, the so called "short-tracks" strategy was proposed in Calamante et al. (2012c). The idea is to impose an upper limit on the tracking distance (*i.e.* when reached, the tracking procedure simply stops, but the track is *not* rejected); thereby directly avoiding the "overcontribution" of tracks to longer tracts. If the maximum track length is set not much higher than the minimum track length, many tracks that make it into the distribution (*i.e.* that achieve the minimum track length) typically also reach the maximum track length. As such, most tracks in the distribution are of equal – or at least very similar – length and per consequence also contribute to a more equal amount of voxels. The resulting track densities end up being distributed more evenly over the brain. Because the shorter tracks overall contribute to a smaller number of voxels, more tracks are needed to retain a good contrast-to-noise ratio (CNR); yet the processing time and amount of storage space required for the final tractogram remain similar (Calamante et al., 2012c).

In Calamante et al. (2012c), the short-tracks strategy was employed to reduce the dynamic range of DEC TDI maps, thereby putting more emphasis on the visualization of the color contrast and the angular information that it encodes.

Because our focus in this work is also on the angular information provided by TODI, we will consistently calculate the TOD based on whole-brain short-tracks tractograms. This should avoid overly large differences in TOD amplitudes between tracts of different length. Do note that TODI can in general of course still be applied to any tractogram resulting from any fiber tracking pipeline; if the aim is for instance to investigate certain properties of such a tractogram.

An example illustrating the difference between a "regular" and a short-tracks distribution is shown in Fig. 6.3. Both distributions were obtained by enforcing a minimum total track length of 15 mm, so it is equally hard for a track to be accepted into either one of them. In the short-tracks case, however, the bidirectional tractography process launched from each seed was limited to a maximum distance of 15 mm in *both* directions; *i.e.* the maximum total track length equals 30 mm. The histogram of the track lengths of the regular track distribution indicates the presence of tracks from a wide range of lengths, of which the longer ones will relatively overcontribute to the (longer) tracts in which they are able to fit. In the short-tracks distribution, on the other hand, the track lengths are limited to a small range of short lengths. Therefore, individual tracks are simply unable to run along the full length of a (longer) tract; hence the track lengths become independent of the tract lengths. The resulting short-tracks (DEC) TDI map (Fig. 6.3) provides a more uniform contrast. Finally, note that it is not unexpected to still see a decent amount of tracks ending up with a length below the absolute maximum of 30 mm: due to the tracks being limited to 15 mm in *both* directions from the seed point, this will even happen in a hypothetical noise-free dataset for any seed point within 15 mm of a tract ending.

Because the short-tracks TOD features distinct lobes of tract length independent amplitude along the general directions of white matter pathways, it bears a qualitative similarity to the FOD from CSD (Tournier et al., 2007) up to some extent in many voxels. However, with regard to such comparisons, it is very important to stress that the TOD and FOD have, by definition and construction, a *completely different meaning*. The amplitude of the FOD along a given orientation can be associated with the intra-axonal volume of (coherently ordered) axons along this orientation; *i.e.* it is mostly sensitive to the partial volume fractions of underlying white matter fiber populations. The FOD's amplitude is in this context referred to as the *apparent fiber density* (AFD) (Raffelt et al., 2012b). The amplitude of a TOD, on the other hand, in general reflects the presence of fiber tracks along certain directions. As mentioned before, fiber tracks – and thus TOD lobes – can only be present at specific locations and orientations if they are supported by the neighborhood, under the tractography constraints. The amplitude of the *short-tracks* TOD specifically, can therefore be interpreted as a measure of *track-like local support* (TLS);

"track-like" due to the enforced tractography constraints, and "local" due to the short-tracks mechanism and the implied tract length independence.

6.2.5 TOD-based tractography and multi-level TODI

The continuity constraints (*e.g.* minimum track length and maximum curvature) to which tracks typically have to adhere, will guarantee a certain degree of mutual consistency between track distributions of neighboring voxels in continuous regions along tracts. Hence, such regional consistency also applies to the TOD lobes associated to these tracts. This ties in perfectly with the interpretation of the short-tracks TOD's amplitude as a measure of TLS: the lobes actually indicate the directions along which support should be present, *i.e.* other TOD lobes along a similar orientation. The TOD lobes thus "connect" nicely with each other up to some extent. These qualities render the short-tracks TOD itself a potentially very interesting candidate for guiding a tractography algorithm: the TOD could as such take on the role of the FOD in existing FOD-based tracking strategies, yielding TOD-based tractography algorithms. The TLS interpretation of the short-tracks TOD's amplitude should allow it to guide a track along directions that are more likely to correspond to *continuous structure over a longer distance* (*i.e. track-like* structure!), because the TOD lobes are effectively oriented along directions where such support originated from in the first place.

We can even build further upon this concept by employing TOD-based tractography itself to generate another whole-brain short-tracks tractogram. From this tractogram, yet another TOD can also be calculated. In order to distinguish between both TODs, we will assign them a "level": the original TOD is linked to "level 1", the other one to "level 2". This system can theoretically be repeated, yielding higher levels. As the FOD is the predecessor of the level 1 TOD in this pipeline, we will assign it "level 0". A schematic overview of the resulting workflow is presented in Fig. 6.4. At level 0, the FOD is typically visualized as an overlay on a DEC FA map. At higher levels, it makes more sense to overlay the TOD on a DEC TDI map calculated from the same tractogram: the combination of both will aid visual assessment of the track distribution. Reasoning the other way around, one could combine the FOD with a map of the total AFD (representing the integral of the FOD over the angular domain, analogously to the relation between the TDI map and the TOD); but since a "DEC AFD" map is not a common concept, we will adhere to the DEC FA map for visualization purposes at "level 0". In this work, we thus explore the differences in quality between the FOD overlaid on the DEC FA map and the level 1 and 2 TOD overlaid on the accompanying DEC TDI map. Finally, the FOD as well as the TOD (levels 1 and 2) will also be used to guide

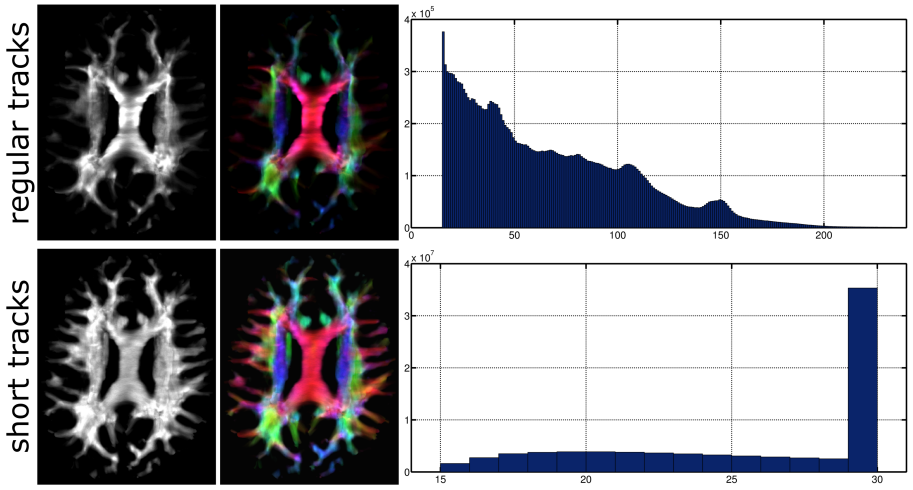

Figure 6.3: Comparison of a whole-brain "regular" track distribution (20 million tracks, min. track length = 15 mm) and short-track distribution (80 million tracks, min. track length = 15 mm, max. distance (bidirectionally) from seed = 15 mm, *i.e.* max. track length = 30 mm). TDI and DEC TDI maps of a slice through the volume are shown, as well as histograms of the track lengths (in mm).

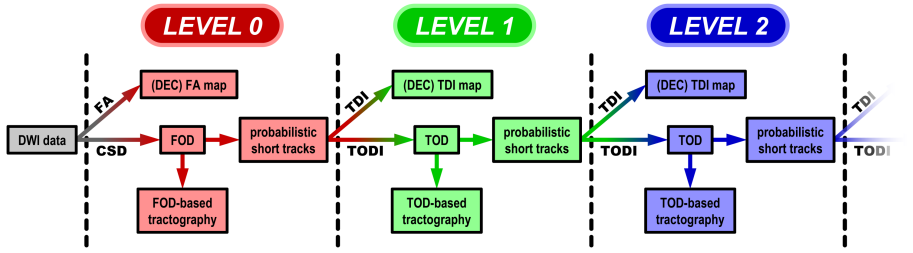

Figure 6.4: Schematic overview of the workflow put forward in this work. Starting from the DWI data, the DEC FA map and FOD are calculated at "level 0". The FOD serves as the basis for (targeted) tractography and is used as well to generate a (whole-brain) probabilistic short-tracks distribution. Starting from the latter, the DEC TDI map and TOD are calculated at "level 1". The process is repeated by employing the short-tracks TOD itself to guide a tractography algorithm.

several (targeted) tractography experiments. The aim of these experiments is twofold: while they directly assess the difference in potential of FOD-based and TOD-based tractography, they can also indirectly provide further insight in the qualitative differences between the FOD and the TOD.

6.2.6 Data

The first dataset consists of an *in vivo* multi-shell acquisition of a healthy human subject. The subject was scanned in a Siemens 3T scanner using a twice-refocused spin-echo (TRSE) sequence (Reese et al., 2003) at a $2.5 \times 2.5 \times 2.5\,\mathrm{mm}^3$ isotropic voxel size. In addition to 10 non-DWI volumes (which were averaged), 3 DWI shells were acquired, each with a different diffusion weighting and a unique set of diffusion weighted directions. The set of directions was independently generated for each shell by electrostatic repulsion (Jones et al., 1999a). In this work, we process the data as if they were 3 separate single-shell datasets of different quality:

- 75 directions, $b = 2800\,\mathrm{s/mm}^2$ (HARDI)
- 40 directions, $b = 1000\,\mathrm{s/mm}^2$ (trade-off)
- 25 directions, $b = 700\,\mathrm{s/mm}^2$ (DTI-like)

Whereas the first shell of HARDI quality is perfectly suited for CSD and subsequent FOD-based tractography (Tournier et al., 2004, 2007, 2013a, 2012), the other shells are expected to provide a bigger challenge.

The second dataset is an *in silico* numerical phantom generated by the *Numerical Fibre Generator* (NFG) software package (Close et al., 2009). We used the pregenerated "*Phantom A*" that was presented in Close et al. (2009) and was also used in Calamante et al. (2011) for the validation of the super-resolution properties of TDI. The phantom represents a sphere that is tightly packed with twisting fiber bundles, including complex fiber architectures such as "kissing", "crossing" and "branching" bundles. The dataset is a volume of $20 \times 20 \times 20$ voxels, which we defined to be of $2.5 \times 2.5 \times 2.5\,\mathrm{mm}^3$ isotropic size (in order to be consistent with our *in vivo* data). The DWI images represent a single-shell acquisition of 60 uniformly distributed DWI-directions at $b = 3000\,\mathrm{s/mm}^2$. We added random Rician noise to the images in order to obtain a signal-to-noise ratio (SNR) of 17 (as was also done in Calamante et al. (2011)).

The third dataset consists of an acquisition of the hardware phantom that was used for the *Fiber Cup* (Fillard et al., 2011); a contest where 10 fiber tracking pipelines were evaluated and compared. More recently, it was also employed as

part of a newly proposed strategy to evaluate tractography pipelines, named *Tractometer* (Côté et al., 2012, 2013). The phantom itself is a disc, containing artificially constructed straight or smoothly bending bundles (Poupon et al., 2008, 2010). At specific locations, it features complex architectures such as "crossing", "splitting" and "kissing" bundles. The phantom was originally scanned in a Siemens 3T scanner, using 6 different acquisition setups with different spatial resolutions and diffusion weightings. As our goal is not to present the best possible full pipeline for the *Fiber Cup* or the *Tractometer*, but rather to directly compare FOD-based and TOD-based tractography under the same conditions, we simply chose one of the acquisitions. The dataset we used, has a $3 \times 3 \times 3\,\mathrm{mm}^3$ isotropic voxel size and a volume of $64 \times 64 \times 3$ voxels. It contains 2 repetitions of DWI images for 64 uniformly distributed DWI-directions at $b = 1500\,\mathrm{s/mm}^2$. This particular acquisition was preferred by most (8 out of 10) participants of the *Fiber Cup* (Fillard et al., 2011) and was also shown to provide the overall best results according to the *Tractometer* (Côté et al., 2013).

6.2.7 Experiments

As a first step, CSD (Tournier et al., 2007) was applied to all individual datasets in order to obtain the FODs. The *in vivo* data was processed as 3 independent single-shell datasets. We consistently performed CSD up to a maximum SH order of 8 for all data. Hence, for the non-HARDI shells of the *in vivo* data, the *super-resolved* version of CSD (Tournier et al., 2007) was required in order to obtain the 45 necessary parameters from the limited set of measurements. Specifically in case of the *Fiber Cup* phantom, the FOD was only reconstructed in the "white matter" (WM) region (Fig. 6.18) and defined to be of zero amplitude elsewhere. All DEC FA maps were calculated from a simple diffusion tensor fit of the data. Visualizations of the FODs will always be presented as an overlay on the DEC FA maps. From this point onwards, all further processing is (indirectly) based on the FODs, as shown in Fig. 6.4.

As mentioned before, we always perform TODI in this work by calculating $\widetilde{\mathrm{TOD}}_{16}(r, u)$ from a dense whole-brain (or phantom) short-tracks tractogram. In the interest of compact notation, we will simply refer to it as the TOD, optionally specifying its spatial resolution and "level". The complete central backbone of the workflow (Fig. 6.4) was calculated at the native resolution of each dataset, up to and including the level 2 TOD. The two dense short-tracks tractograms along this pipeline were obtained from probabilistic FOD- and TOD-based tracking. Both employ the same parameters, apart from the FOD and TOD thresholds: the latter were chosen to reflect the difference in order of magnitude between the FODs and TODs for each dataset. Tracking is

always performed bidirectionally from each seed point. In order to obtain the aforementioned *short-tracks* distribution, the maximum distance that tracks are allowed to travel in *each of both* directions is always set to the same value as the minimum *total* track length. As such, the maximum *total* track length equals two times the minimum *total* track length. The other parameters were set as follows:

- *in vivo* datasets: 80 million tracks; min. FOD (TOD) thresholds to initiate/continue tracks are 0.2 (10000) / 0.1 (5000); step size is 0.2 mm; min. radius of curvature is 1 mm; min. total track length is 15 mm.
- *in silico* phantom: 80 million tracks; min. FOD (TOD) thresholds to initiate/continue tracks are 0.2 (20000) / 0.1 (10000); step size is 0.2 mm; min. radius of curvature is 1 mm; min. total track length is 5 mm.
- *Fiber Cup* phantom: 80 million tracks; min. FOD (TOD) thresholds to initiate/continue tracks are 0.2 (200000) / 0.1 (100000); step size is 0.2 mm; min. radius of curvature is 1 mm; min. total track length is 24 mm.

Based on the resulting tractograms, the level 1 and level 2 DEC TDI maps were also obtained at the native resolutions, to complement the visualization of the level 1 and level 2 TODs. Additionally, for all *in vivo* datasets, we generated a super-resolution version of the level 1 and level 2 DEC TDI maps at a $0.5 \times 0.5 \times 0.5\,\text{mm}^3$ isotropic voxel size. Finally, specifically for the *in vivo* HARDI dataset, we also calculated such a super-resolution version as well as an extreme sub-resolution version (a single voxel encapsulating the whole short-tracks distribution) of the level 1 TOD.

Different (targeted) tractography experiments were performed for each dataset, based on the FOD as well as the level 1 and level 2 TOD at the native resolution. For these experiments, no maximum track length is enforced: tracks are always allowed to continue (bidirectionally) until the minimum FOD or TOD threshold is no longer met. If a track has not reached the minimum track length at that point, it is rejected. Tracks are generated until the required amount is successfully accepted. Specifically for deterministic tractography, there is also no curvature constraint in place: the tracks always follow the local FOD or TOD maximum closest to their current direction. The experiments and further parameters for each dataset are defined as follows:

- *in vivo* datasets: deterministic and probabilistic tractography; tracks started from 3 seed regions (genu of the corpus callosum (GCC), cingulum bundle (CB), superior cerebellar peduncle (SCP)) along a predetermined initial direction with an angular tolerance of 30 ° (Fig. 6.13); 10000

tracks required for each seed region; min. FOD (TOD) thresholds to initiate/continue tracks are 0.2 (10000) / 0.1 (5000); step size is 0.2 mm; min. radius of curvature is 1 mm (only for probabilistic tracking); min. total track length is 15 mm.

- *in silico* phantom: only deterministic tractography; tracks started from 3 spherical seed regions with a radius of 1 mm (Fig. 6.17D); 10000 tracks required for each seed region; min. FOD (TOD) thresholds to initiate/continue tracks are 0.2 (20000) / 0.1 (10000); step size is 0.2 mm; min. total track length is 5 mm.
- *Fiber Cup* phantom: deterministic and probabilistic tractography; tracks started from the whole WM region (Fig. 6.18); 1000000 tracks required; min. FOD (TOD) thresholds to initiate/continue tracks are 0.2 (200000) / 0.1 (100000); step size is 0.2 mm; min. radius of curvature is 1 mm (only for probabilistic tracking); min. total track length is 24 mm.

Apart from visualizing the (qualitative) results, we will also report on the total amount of tracks that had to be generated for each experiment in order to obtain the required amount of accepted tracks. As stated before, the excess tracks were rejected because they failed to even reach the minimum track length.

Finally, specifically for the tractograms resulting from the experiments on the *Fiber Cup* phantom, we performed a quantitative validation in the spirit of the recently proposed *Tractometer* (Côté et al., 2012, 2013). By adopting some of the terminology that was introduced in Côté et al. (2013), we subdivided the 1000000 tracks of each tractogram in three categories:

- Valid Connections (VC): tracks connecting one of the 7 valid pairs of regions along the associated topologically valid path, as indicated in Fig. 6.18.
- Invalid Connections (IC): tracks connecting either an invalid pair of regions, or a valid pair of regions along a topologically invalid path (*e.g.* connecting both regions labeled "3" along a path through the crossing region indicated by the yellow box in Fig. 6.18).
- No Connections (NC): tracks failing to connect a pair of regions.

The amount of tracks in each category is reported in %; since each track is assigned to exactly one category, they all sum up to 100 %. We also report on the amount of VC /(VC + IC). This value represents the amount of VC in case anatomically-constrained tractography (ACT) priors (Smith et al., 2012) would be enforced; *i.e.* if tracks must connect gray matter (GM) or otherwise relevant

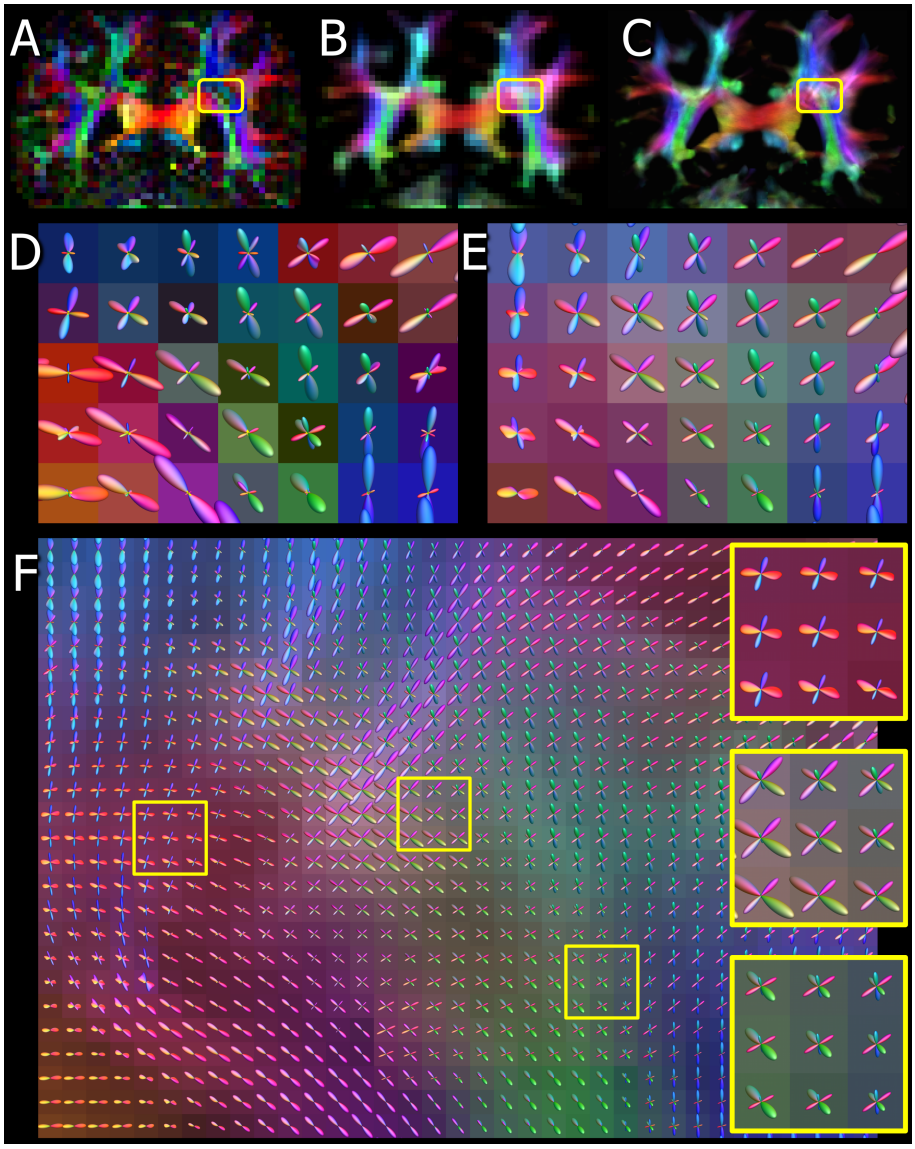

Figure 6.5: TODI at native resolution and super-resolution, in a coronal slice of the *in vivo* HARDI dataset. (A) DEC FA map. (B) DEC TDI map at native resolution and (C) super-resolution. (D–F) Zoomed visualization of the boxed region indicated in (A–C), with an additional overlay of (D) the FOD, (E) the TOD at native resolution and (F) super-resolution. Further zoomed regions are also presented in (F).

regions to be accepted, and therefore NC are rejected. As is evident from the above definitions, all these *Tractometer* measures focus on a quantitative assesment of the global *connectivity* aspect, and as such nicely complement more local or qualitative validation.

All processing and experiments were performed using *MRtrix* (Tournier et al., 2012). The proposed TODI method was implemented by adapting the existing *MRtrix* functionality for calculating (DEC) TDI maps.

6.3 Results

6.3.1 Native and super-resolution TODI

In Fig. 6.5, the resulting (level 1) TOD of the *in vivo* HARDI dataset is shown at native resolution ($2.5 \times 2.5 \times 2.5\,\mathrm{mm}^3$) and super-resolution ($0.5 \times 0.5 \times 0.5\,\mathrm{mm}^3$). It can be appreciated from the comparison between the DEC FA map (Fig. 6.5A) and the DEC TDI map (Fig. 6.5B) that the color contrast of the latter is already "better behaved". Because a color in the DEC FA map can only partially encode a single direction and is directly based on the orientation of the principal eigenvector of the tensor, it becomes ill-defined in voxels containing complex fiber architecture such as the typical fiber crossings. Combined with the fact that FA is typically reduced in these regions, the DEC FA map offers little expressive power for such complex fiber configurations. The DEC TDI map, on the other hand, *sums* the contributions (colors) of the different track populations present in the voxel. Therefore, its color definition and overall appearance is much more robust: in areas where different tracts meet, their local colors are intuitively mixed. The colors also show a better correspondence with their surroundings, due to the continuity constraints on the underlying short-tracks distribution. However, the final color in each voxel still lacks the expressive power to even represent the single average direction. Furthermore, in areas such as the boxed region indicated in Fig. 6.5A–C, the configuration gets so complex that, even for a more global view of such a region, the DEC TDI map is limited in its visual expression. The TOD, on the other hand, already yields an accurate description of the track distribution at the native resolution (Fig. 6.5E). As a direct consequence of the earlier mentioned continuity constraints on the short-tracks distribution, it also shows increased spatial and angular coherence as compared to the FOD (Fig. 6.5D). The super-resolution TOD (Fig. 6.5F) adds upon this by allowing for easier "visual tracking" of several track populations throughout the region. Provided enough tracks are generated, the angular configuration of the tracks at the super-resolution voxel level is still accurately described. The insets in Fig. 6.5F show some of these very local angular

distributions. Based only on their underlying (and even surrounding) colors, it is impossible to recover these configurations.

Finally, we should note that the *absolute* amplitude of the complete (5D) FOD or TOD in each visualization (such as Fig. 6.5D–F) is always scaled by a single constant factor to overcome the difference in order of magnitude between FOD and TOD. Also, as mentioned before, the TOD amplitude represents a measure of TLS and therefore clearly differs in interpretation compared to the FOD amplitude (AFD). This explains the differences in *relative* amplitude among FOD lobes (Fig. 6.5D) and TOD lobes (Fig. 6.5E); both bear unique meanings and thus yield unique contrasts and information.

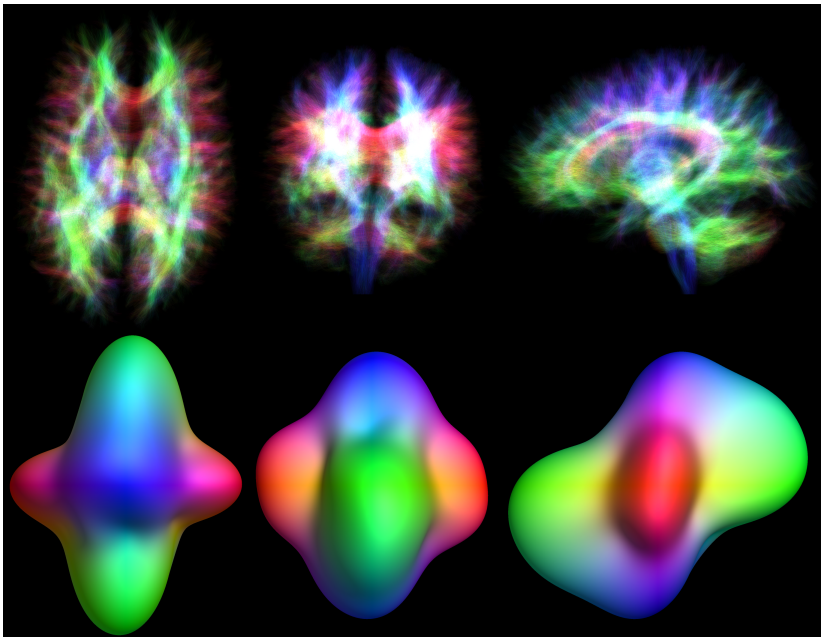

Figure 6.6: TODI at extreme sub-resolution: a single voxel encapsulating the whole-brain short-tracks distribution of the *in vivo* HARDI dataset. *Top row:* axial, coronal and sagittal projection of a subset (0.5 million) of all (80 million) short-tracks. *Bottom row:* the single voxel TOD of all (80 million) short-tracks; as viewed from the bottom, front and left side.

6.3.2 Extreme sub-resolution TODI

An extreme sub-resolution version of the (level 1) TOD, where the whole-brain short-tracks tractogram of the *in vivo* HARDI dataset is encapsulated in a single voxel, is shown in Fig. 6.6 from three perpendicular viewing angles (bottom, front, left) alongside corresponding projections (axial, coronal, sagittal) of a subset of all short-tracks. This TOD as such presents a global view on the orientational organization of the short-tracks distribution, which gives the impression of being dominated by 3 main angular components. A clear structure consisting of dominant mediolateral, anteroposterior and superoinferior components can be observed. While the mediolateral component has a perpendicular orientation relative to the plane of the other components (*i.e.* the midsagittal plane), the anteroposterior and superoinferior components are however not mutually perpendicular (but can still be discerned qualitatively). The mediolateral component also showns more dispersion in a coronal plane that is specifically oriented along the superoinferior component. The accurate mapping of this angular distribution was only possible through our practice of fully integrating the directional information along each track.

6.3.3 Multi-level TODI and TOD-based tractography

In vivo datasets

An overview of level 0 DEC FA maps compared to level 1 and level 2 DEC TDI maps in axial, coronal and sagittal slices, is presented for the *in vivo* HARDI (Fig. 6.7), trade-off (Fig. 6.9) and DTI-like (Fig. 6.11) datasets. The corresponding FODs (level 0) and TODs (levels 1 and 2) for the indicated boxed regions are shown in Fig. 6.8 (HARDI), Fig. 6.10 (trade-off) and Fig. 6.12 (DTI-like). This region in the centrum semiovale is known to contain three major crossing populations of commissural, association and projection fibers. The FODs resulting from the trade-off and DTI-like datasets show a lower quality compared to the FODs obtained from the HARDI dataset. The reason for this is twofold. Due to a lower b-value (in the trade-off and DTI-like datasets), the angular contrast is reduced, and thus CSD will yield less robust results. Furthermore, fewer DWI-directions were sampled for these lower b-value datasets. Hence, the angular resolution is also reduced. Our aim here was to compare how TODI performs on datasets (and the resulting FODs) of such different qualities. The lower quality datasets pose a greater challenge: the orientations (and even amplitudes) of the FOD lobes in each voxel are less consistent in relation to their neighborhood, and the FODs show a good amount of spurious peaks as well. As compared to the FOD, the level 1 TOD in each voxel shows increased

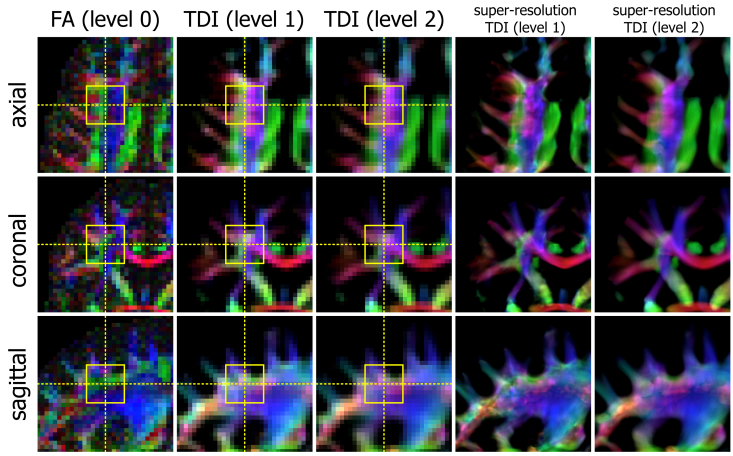

Figure 6.7: Level 0 DEC FA map compared to level 1 and level 2 DEC TDI maps (native resolution and super-resolution) in an axial, coronal and sagittal slice of the *in vivo* HARDI dataset (75 directions, $b = 2800\,\mathrm{s/mm^2}$). The dashed lines indicate the position of the other slices. The boxed regions are shown in detail in Fig. 6.8.

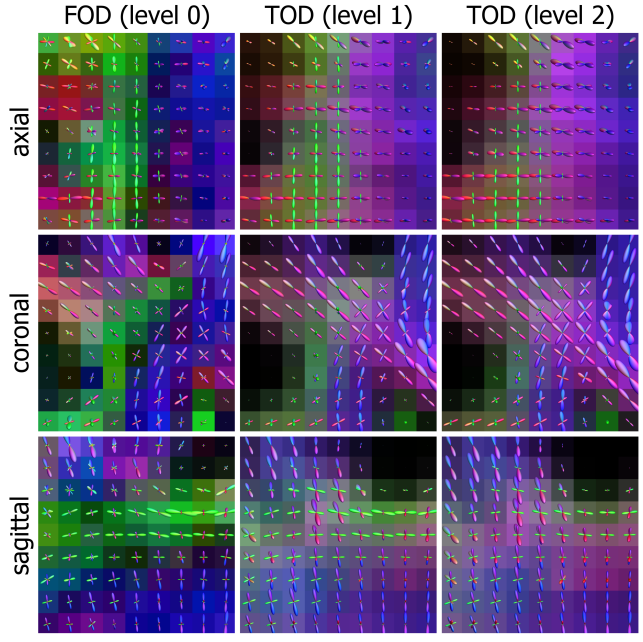

Figure 6.8: Level 0 FOD compared to level 1 and level 2 TOD (native resolution) of the *in vivo* HARDI dataset, visualized in the boxed regions of Fig. 6.7.

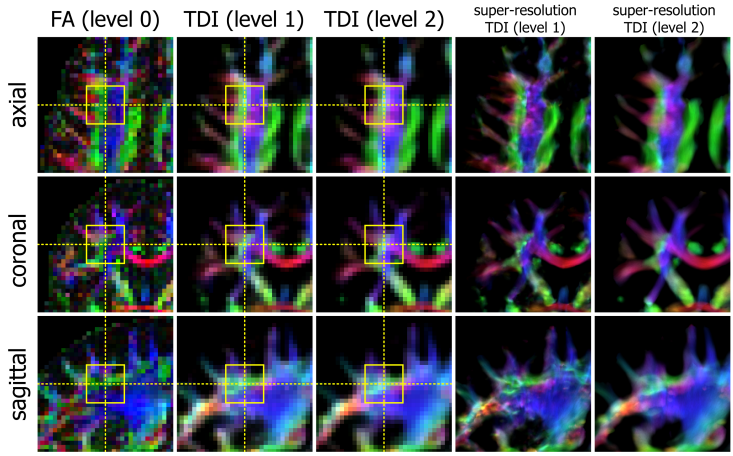

Figure 6.9: Level 0 DEC FA map compared to level 1 and level 2 DEC TDI maps (native resolution and super-resolution), in an axial, coronal and sagittal slice of the *in vivo* trade-off dataset (40 directions, $b = 1000\,\mathrm{s/mm^2}$). The dashed lines indicate the position of the other slices. The boxed regions are shown in detail in Fig. 6.10.

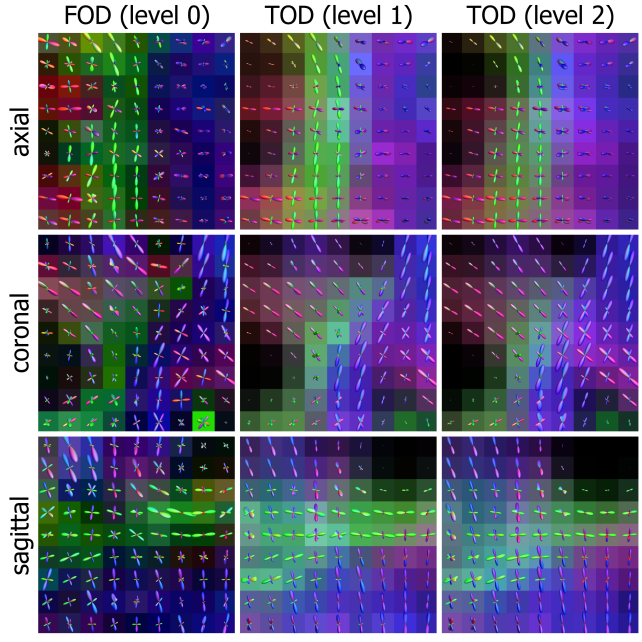

Figure 6.10: Level 0 FOD compared to level 1 and level 2 TOD (native resolution) of the *in vivo* trade-off dataset, visualized in the boxed regions of Fig. 6.9.

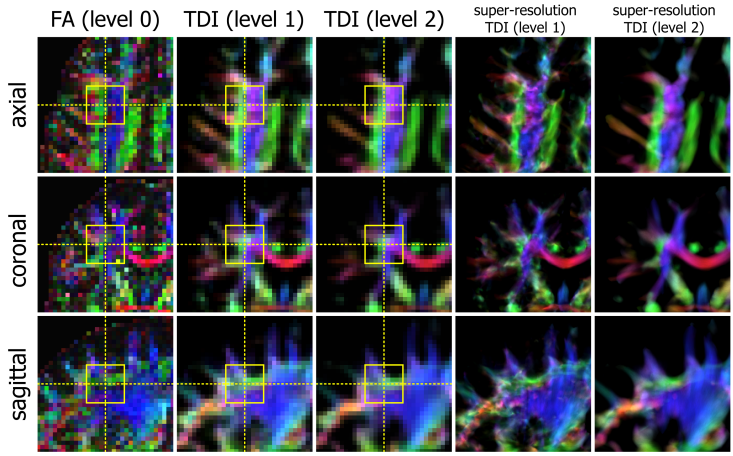

Figure 6.11: Level 0 DEC FA map compared to level 1 and level 2 DEC TDI maps (native resolution and super-resolution), in an axial, coronal and sagittal slice of the *in vivo* DTI-like dataset (25 directions, $b = 700\,\mathrm{s/mm^2}$). The dashed lines indicate the position of the other slices. The boxed regions are shown in detail in Fig. 6.12.

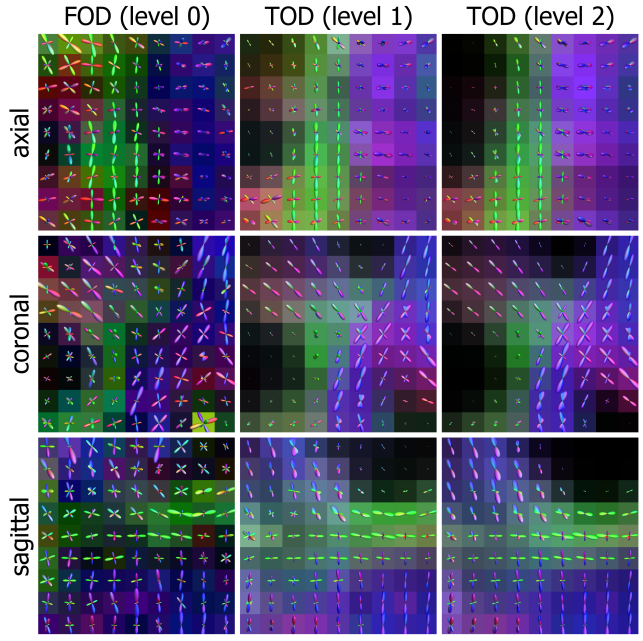

Figure 6.12: Level 0 FOD compared to level 1 and level 2 TOD (native resolution) of the *in vivo* DTI-like dataset, visualized in the boxed regions of Fig. 6.11.

spatial and angular consistency with its surroundings due to the continuity constraints that were imposed on the short-tracks distribution; this is also in line with the TLS interpretation of the amplitude. The benefits can be seen in all datasets, but they are most striking in those of lower quality; *i.e.* where there still exists a large margin for improvement in terms of consistency. A clear example of this is the coronal view of *e.g.* the DTI-like dataset (Fig. 6.12): while the crossing area between the commissural and projection fibers seems heavily "disrupted" in the FOD, it is almost fully recovered in the level 1 TOD. Spurious peaks are also eliminated because they are simply not (or barely) supported by their surroundings during the fiber tractography. Further improvement can be seen in the level 2 TOD, albeit more subtle. This can be attributed to the fact that there is less room for improvement after a first application of TODI. While the differences between level 1 and level 2 are limited at the native spatial resolution, they are a lot easier to characterize from the super-resolution DEC TDI maps provided in Fig. 6.7, Fig. 6.9 and Fig. 6.11: the images at level 2 are less affected by local perturbations (and in fact most of the highly variable texture and features that could induce overinterpretation, as we warned for in Dhollander et al. (2012)). These differences are again most apparent in the datasets of lower quality, *i.e.* in the most challenging cases. Note that each of these super-resolution DEC TDI maps actually represents a detailed visualization of a (probabilistic short-tracks) tractography experiment. The level 1 and level 2 super-resolution DEC TDI maps as such already yield a limited view on some of the qualitative differences between FOD-based and (level 1) TOD-based probabilistic tracking outcomes. Finally, the coronal view of *e.g.* the HARDI dataset (Fig. 6.8) provides another nice example of the differences in relative amplitude among FOD lobes and TOD lobes: within the crossing region, the commissural fiber population shows a lower AFD in comparison to the projection fiber population, but at the same time a higher amount of TLS. This shows again that both AFD and TLS are unique measures, and the short-tracks TOD is not just a regularized version of the FOD.

The actual results of the specific deterministic and probabilistic targeted tractography experiments, seeded from the regions indicated in Fig. 6.13 (GCC, CB, SCP), and based on the FOD as compared to the TOD (levels 1 and 2) are presented for the *in vivo* HARDI (Fig. 6.14), trade-off (Fig. 6.15) and DTI-like (Fig. 6.16) datasets. The FOD-based tracking results are better for the HARDI dataset as compared to the trade-off and DTI-like datasets: fewer spurious tracks (false positives) and less missing structures (false negatives) are present, as well as a more continuous appearance of the results. This is fully in line with the previously explained findings on the quality of the FODs for these different datasets. In general, probabilistic tracking leads to a higher amount of dispersion in the distribution of the tracks and can as such provide a more complete view of several structures (on the other hand, more

false positives could potentially be included). The FOD-based deterministic tractography, for instance, has great difficulties in reconstructing the fanning structure of the projection fibers in the corona radiata when seeding from the SCP and tracking upwards, even for the HARDI dataset (Fig. 6.14); while its FOD-based probabilistic counterpart consistently performs better at this task. Another challenge for both deterministic and probabilistic tracking is the presence of sharp turns, such as the one at the back of the CB. Especially for the lower quality datasets, this proves to be a difficult task. The level 1 TOD-based tractography (deterministic as well as probabilistic) resolves a lot of these problems. Furthermore, fewer false positive/negative tracks and more coherent, clearly delineated bundles of tracks can be observed. Level 2 TOD-based tractography again improves upon this, albeit in a limited amount and mostly in those extra challenging situations (*e.g.* sharp bending and fanning in the lower quality datasets). We should however also mention that TOD-based tracking can still run into problems when the data quality is very limited. One such case is the deterministic tractography seeded from the SCP in the DTI-like dataset (Fig. 6.16), where a large frontal part of the tracks is missing. The probabilistic TOD-based tracking performs better, but still shows difficulties in this specific case. A better strategy to recover such a structure would consist of seeding throughout the whole brain and retaining those tracks that pass through the SCP region (instead of only seeding from this region). In this case, however, we were merely interested in a direct comparison of FOD-based and TOD-based tractography under the same simple and clearly defined conditions, rather than the best possible final result. Note that each individual tractogram

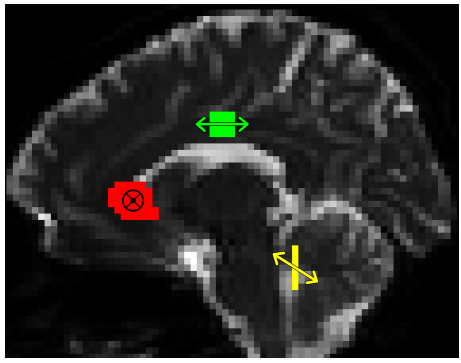

Figure 6.13: Sagittal projection (on a slice of the average non-DWI ($b = 0$) volume) of the seed regions used for targeted tractography in the *in vivo* datasets: the genu of the corpus callosum (GCC) (*red*), cingulum bundle (CB) (*green*) and superior cerebellar peduncle (SCP) (*yellow*). For each experiment, tracks are seeded uniformly in one of these regions along the indicated initial direction (with an angular tolerance of 30 °).

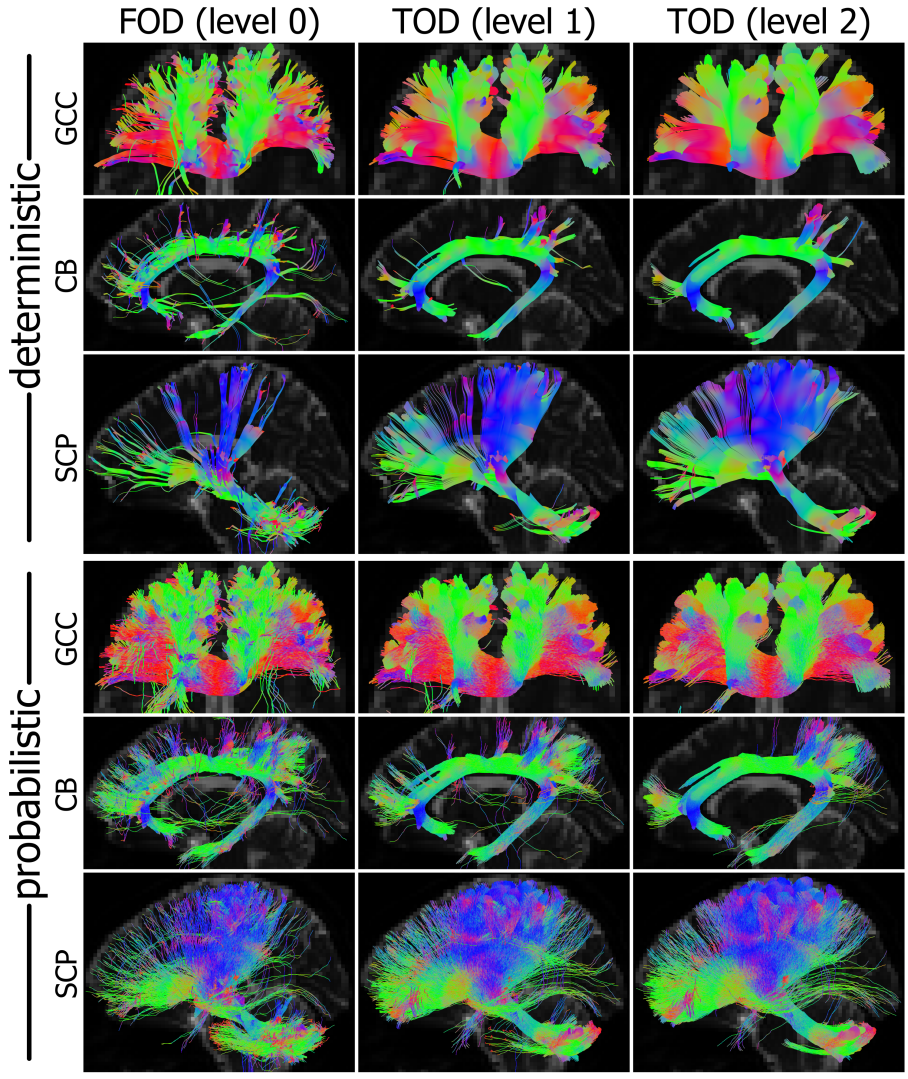

Figure 6.14: Targeted tractography results for the *in vivo* HARDI dataset (75 directions, $b = 2800\,\mathrm{s/mm^2}$), each consisting of 10000 tracks seeded uniformly from the regions in Fig. 6.13. Deterministic as well as probabilistic FOD-based (level 0) and TOD-based (levels 1 and 2) tractography results are compared.

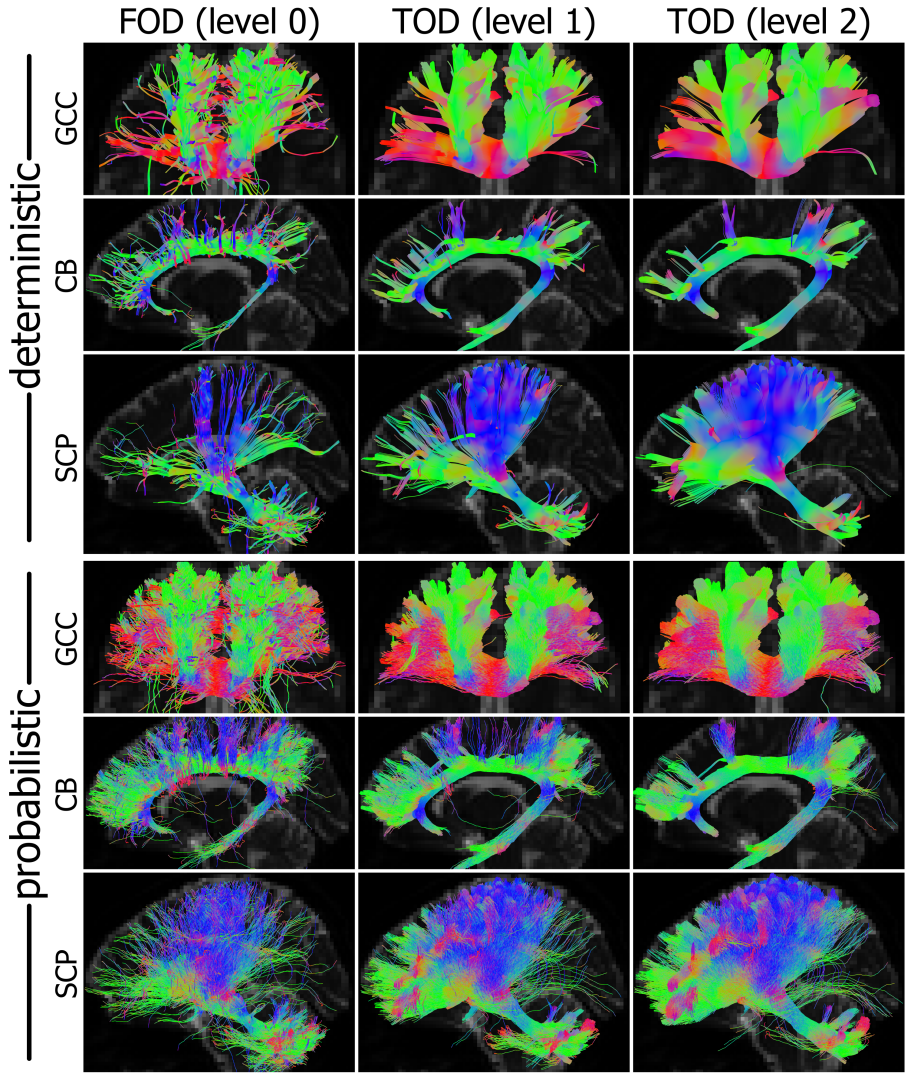

Figure 6.15: Targeted tractography results for the *in vivo* trade-off dataset (40 directions, $b = 1000\,\mathrm{s/mm^2}$), each consisting of 10000 tracks seeded uniformly from the regions in Fig. 6.13. Deterministic as well as probabilistic FOD-based (level 0) and TOD-based (levels 1 and 2) tractography results are compared.

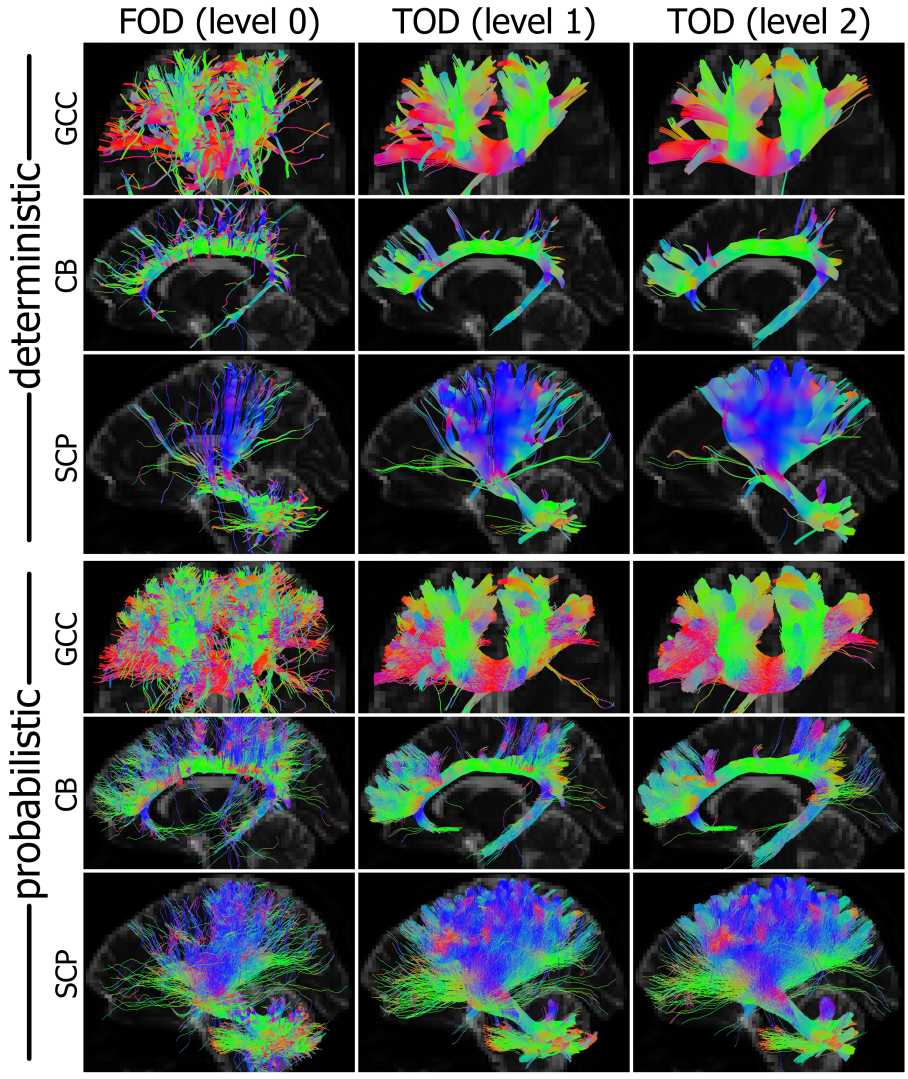

Figure 6.16: Targeted tractography results for the *in vivo* DTI-like dataset (25 directions, $b = 700\,\mathrm{s/mm^2}$), each consisting of 10000 tracks seeded uniformly from the regions in Fig. 6.13. Deterministic as well as probabilistic FOD-based (level 0) and TOD-based (levels 1 and 2) tractography results are compared.

Table 6.1: Total amount of tracks that had to be generated for each tractography experiment on the *in vivo* datasets, in order to obtain 10000 accepted tracks. The excess tracks were rejected because they failed to reach the minimum track length (15 mm) while maintaining the other constraints.

Deterministic	HARDI			trade-off			DTI-like		
	FOD 0	TOD 1	TOD 2	FOD 0	TOD 1	TOD 2	FOD 0	TOD 1	TOD 2
GCC	10106	10020	10012	10303	10014	10011	10845	10049	10020
CB	10243	10006	10000	10770	10005	10000	11401	10037	10001
SCP	10579	10021	10006	11150	10018	10001	10630	10019	10000
Probabilistic	HARDI			trade-off			DTI-like		
	FOD 0	TOD 1	TOD 2	FOD 0	TOD 1	TOD 2	FOD 0	TOD 1	TOD 2
GCC	10027	10005	10000	10059	10004	10004	10305	10008	10013
CB	10122	10003	10000	10330	10002	10000	10475	10007	10003
SCP	10192	10004	10000	10191	10000	10006	10242	10001	10001

as presented in Fig. 6.14, Fig. 6.15 and Fig. 6.16 consists of exactly 10000 accepted tracks. Tracks that failed to reach the minimum track length (15 mm) were already rejected during the tractography process. For each experiment, the total amount of tracks that had to be generated in order to obtain 10000 tracks (*i.e.* 10000 plus the number of rejected tracks) is reported in Table 6.1. Overall, the amount of rejected tracks is not very high; most probably due to the specific seed regions in combination with the well-chosen (anatomically informed) initial directions (Fig. 6.13). However, the TOD-based tractography still consistently succeeded to reduce the number of rejected tracks by several orders of magnitude as compared to the FOD-based results. All other findings on the relative amounts of rejected tracks are fully in line with the previous findings on the quality of the FOD/TOD and the qualitative tractography results of the accepted tracks: better qualitative findings typically match a lower amount of rejected tracks in Table 6.1. This also means that, if we would not have constrained our tractography experiments to achieve a minimum track length (*i.e.* no tracks would have been rejected), all previous qualitative findings on the differences between the resulting tractograms would have been even more contrasting.

In silico phantom

A slice of the level 0 FOD is compared to the level 1 and level 2 TOD in Fig. 6.17A–C, for the *in silico* phantom. The FOD in all voxels is affected by the noise that was added to the DWI data, resulting in several spurious peaks and deviations of the main lobes' orientations. Voxels in crossing regions are

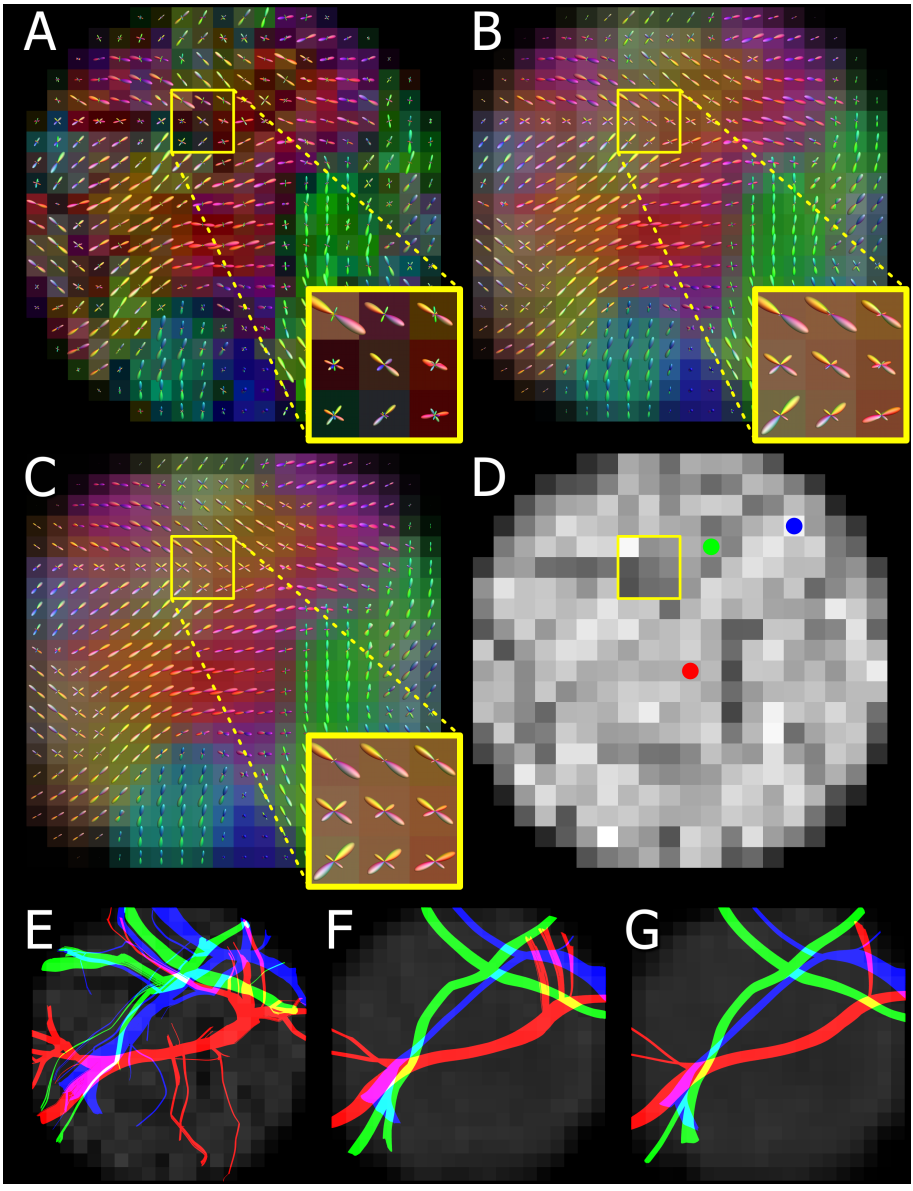

Figure 6.17: TODI results for the *in silico* phantom. (A) Level 0 FOD and DEC FA map. (B) Level 1 and (C) level 2 TOD and DEC TDI map. (D) Location of 3 spherical seed regions (radius = 1 mm). (E–G) "Axial" projections of 3×10000 targeted tracks, obtained by (E) level 0 FOD-based and (F) level 1 and (G) level 2 TOD-based deterministic tractography, seeded uniformly from the regions in (D).

Table 6.2: Total amount of tracks that had to be generated for each tractography experiment on the *in silico* phantom, in order to obtain 10000 accepted tracks. The excess tracks were rejected because they failed to reach the minimum track length (5 mm) while maintaining the other constraints.

	FOD (level 0)	TOD (level 1)	TOD (level 2)
red	20212	10000	10000
green	12720	10000	10000
blue	15684	10000	10000

among the most vulnerable, as they feature lower amplitudes of the individual FOD lobes (each voxel has a similar total amount of AFD; yet in crossing regions it is split up between the different bundles' lobes). In the level 1 TOD, spurious peaks are almost consistently absent, due to a lack of TLS from the surroundings along such directions. Additionally, the orientations of the main TOD lobes for all voxels (in particular those with crossing fibers) show increased regional coherence. Further improvements in the level 2 TOD are again very limited at the phantom's native spatial resolution, and apply mostly to the crossing fiber regions. One such region is pictured in more detail in the zoomed inset (Fig. 6.17A–C).

The FOD-based (level 0) and TOD-based (levels 1 and 2) deterministic tractography results for the *in silico* phantom are shown in Fig. 6.17E–G. The colors correspond to those of the seed regions in Fig. 6.17D. Note that, while the tracks run in full 3D space, the results are simply projected to a 2D slice for the purpose of easy assessment (the seed regions are, however, all centered on the actual slice that is shown in Fig. 6.17A–D). The FOD-based tracking resulted in several erratic and spurious tracks, due to deviations of the main lobes' orientations as well as spurious peaks in the FOD. The level 1 TOD-based tractography results show much smoother and coherent bundles of tracks, corresponding to the structure of this particular phantom. Just as before, the level 2 TOD-based tracking only provided little extra benefit: a slight increase in smoothness is observed and a few (possibly) spurious tracks are removed, but the main bundles' delineated trajectories remain mostly the same. Similarly to the *in vivo* experiments, we required 10000 tracks to be accepted for each seed region by reaching a minimum track length (only 5 mm in this case). The total amount of tracks that were generated to achieve this, is reported in Table 6.2 for the individual *in silico* experiments. The result is quite remarkable: while a decent number of tracks were rejected for each FOD-based experiment (up to more than 50 % for the red seed region!), not a single one was rejected for all of the TOD-based (levels 1 and 2) experiments. In fact, we even noticed that every individual TOD-based track always ran

along the full length of the bundle; *i.e.* consistently reaching the outer surface of the spherical phantom in both directions from the seed region. The rather large amounts of rejected tracks for the FOD-based experiments can probably be attributed to the fact that no initial tracking direction was given for the seed regions. Since the FOD features several spurious peaks, many tracks might have started running along such spurious orientations, subsequently not reaching the required 5 mm because those orientations do not correspond to a continuous bundle. The TOD, on the other hand, already incorporates the TLS information and can immediately guide the tracks in a sensible direction corresponding to a true bundle. Finally also note that, if we had not imposed a minimum track length for these experiments, the qualitative results for the TOD-based tractography (Fig. 6.17F–G) would appear the same, while the FOD-based results (Fig. 6.17E) would include spurious tracks shorter than 5 mm at the cost of longer tracks (proportional to the numbers provided in Table 6.2).

Fiber Cup phantom

The level 0 FOD is compared to the level 1 and level 2 TOD in Fig. 6.19 (*top row*), for a crossing region of the *Fiber Cup* phantom, as indicated by the yellow box in Fig. 6.18. Because the phantom consists of simple synthetic fibers, only an "extra-cellular" hindered diffusion phenomenon is simulated. This resulted in much lower FA values than those typically seen in (human) brain WM regions (Fillard et al., 2011), and an associated severely reduced CNR. The quality of the FOD, especially in crossing regions such as the one depicted in Fig. 6.19 (*top row*), is therefore limited. In line with the findings from the other datasets, the level 1 TOD shows a clear and consistent pattern of TLS that corresponds much better to the underlying perpendicular crossing structure. The level 2 TOD provides again a largely similar outcome; as compared to the level 1 TOD, a slight increase of TLS is observed within the crossing voxels relative to the other voxels. Note that, even though the amplitudes of both the level 1 and level 2 TOD are measures of TLS, they can still show different contrasts; this all depends on the parameters and input of the preceding short-tracks tractography. In all our experiments, however, the parameters to generate the two intermediate short-tracks distributions were always the same for both of them (apart from the FOD/TOD thresholds, but those were chosen to reflect the difference in order of magnitude). The only real cause of possible differences between both (in our current setting) is the input: the first short-tracks distribution is obtained from the FOD, while the second one results from the level 1 TOD.

A subset of the (1000000) deterministic tracks, seeded from the whole "WM" region as indicated in Fig. 6.18, and obtained by FOD-based and (levels 1 and 2) TOD-based deterministic tractography is presented in Fig. 6.19 (*bottom row*)

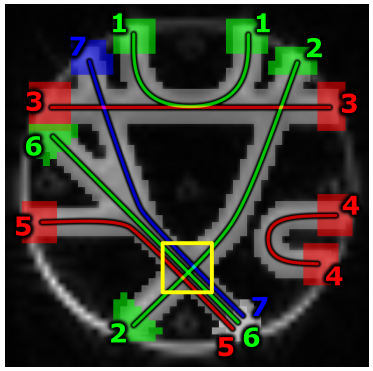

Figure 6.18: Overview of the *Fiber Cup* phantom. The "white matter" (WM) region is overlaid in gray on the middle slice of the average non-DWI ($b = 0$) volume. Each matching pair of numbered regions and the associated path demonstrates one of the 7 ways to establish a topologically valid connection. Qualitative results for the yellow boxed region are shown in Fig. 6.19.

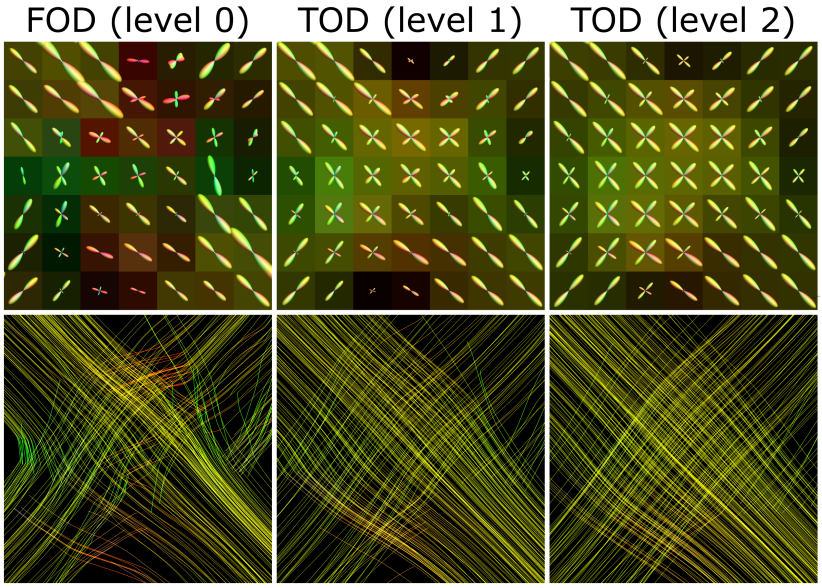

Figure 6.19: TODI results for the *Fiber Cup* phantom, visualized in the boxed region of Fig. 6.18. *Top row:* Level 0 FOD (and DEC FA map), compared to level 1 and level 2 TOD (and DEC TDI map). *Bottom row:* "Axial" projections of subsets of all (1000000) deterministic tracks, obtained by FOD-based and TOD-based (levels 1 and 2) deterministic tractography, seeded uniformly from the WM region (depicted in Fig. 6.18).

Table 6.3: Total amount of tracks that had to be generated for each tractography experiment on the *Fiber Cup* phantom, in order to obtain 1000000 accepted tracks. The excess tracks were rejected because they failed to reach the minimum track length (24 mm) while maintaining the other constraints.

	FOD (level 0)	TOD (level 1)	TOD (level 2)
Deterministic	1140993	1001186	1000585
Probabilistic	1104646	1000677	1000212

for the same crossing region of the *Fiber Cup* phantom. Again, the tracks clearly reflect the quality of the FOD/TOD by which they were guided. The FOD-based results show high sensitivity to the local deviations of the main FOD lobes' orientations within the crossing region. As a result, a lot of tracks were not able to successfully run across this region, or even erroneously connected both different bundles. The much more consistent level 1 TOD managed to guide the tracks reliably by indicating the major orientations of TLS. The most coherent straightly crossing bundles were obtained from the level 2 TOD; but the quality does not differ all that much from the level 1 TOD-based results. The actual experiments consisted of deterministic as well as probabilistic FOD/TOD-based tractography, and each required 1000000 tracks to be accepted for the whole phantom by achieving a minimum track length of 24 mm. The total amount of generated tracks to obtain 1000000 accepted tracks for each experiment, is reported in Table 6.3. In line with all previous findings up to this point, TOD-based tractography resulted in a reduction of the amount of rejected tracks by several orders of magnitude, as compared to FOD-based tractography.

The quantitative *Tractometer* results for the *Fiber Cup* phantom are provided in Table 6.4. Note that even the FOD-based tractography appears to perform much better as compared to the results reported in Côté et al. (2013). This is due to several different choices that were made across the pipeline; *e.g.* a better delineated "WM" region, slightly wider regions to evaluate the connections, a minimum track length (*i.e.* only the 1000000 accepted tracks are taken into account), etc. Our aim was to make reasonable choices that would target a good – but not necessarily the best possible – result. As these choices should equally benefit or otherwise impact the FOD-based as well as the TOD-based results, they still allow a perfectly honest comparison between them. Just as for all other findings, the differences between the level 0 FOD and the level 1 TOD are quite impressive, while those between the level 1 and level 2 TOD are smaller. The TOD-based tractography consistently performs better: an increase of VC and a decrease of IC and NC is observed. By what we already know of the TLS meaning and qualities of the TOD, the decrease of NC comes as no surprise: tracks are able to run across greater lengths along continuous

Table 6.4: Quantitative *Tractometer* results: the amount of valid connections (VC), invalid connections (IC) and no connections (NC) for each tractography experiment on the *Fiber Cup* phantom, reported in % (of the 1000000 accepted tracks). The amount of VC/(VC + IC) is provided as well (in %). The filtered VC for each case are visualized in Fig. 6.20.

Deterministic			
	FOD (level 0)	TOD (level 1)	TOD (level 2)
VC	40.4291	69.1631	75.7147
IC	13.2397	5.8644	4.0825
NC	46.3312	24.9725	20.2028
$\frac{VC}{VC+IC}$	75.3307	92.1837	94.8839
Probabilistic			
	FOD (level 0)	TOD (level 1)	TOD (level 2)
VC	29.3934	50.2976	57.8701
IC	13.9695	13.7285	13.3753
NC	56.6371	35.9739	28.7546
$\frac{VC}{VC+IC}$	67.7847	78.5580	81.2264

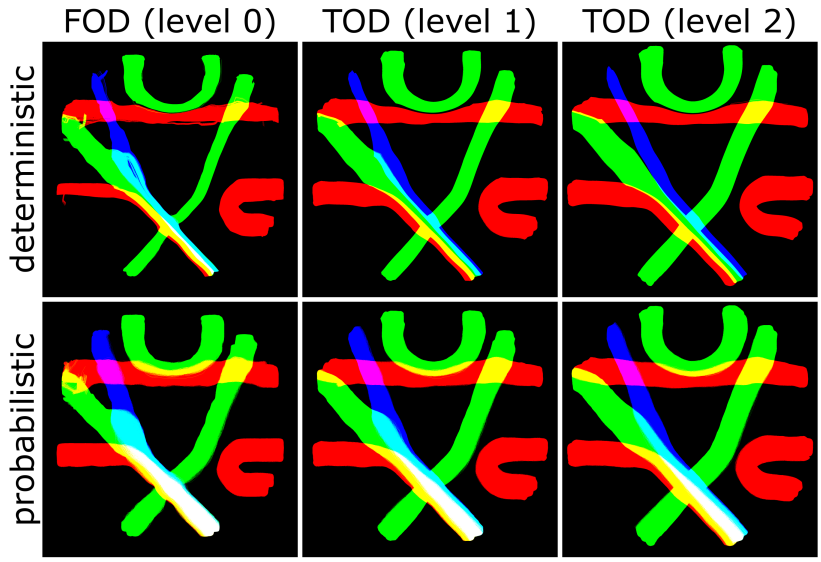

Figure 6.20: Qualitative *Tractometer* results: "axial" projections of the valid connections (VC), filtered from the 1000000 accepted tracks, for each tractography experiment on the *Fiber Cup* phantom. The amount of VC for each case is given in Table 6.4.

bundles and will as such be more likely to actually connect a pair of regions. Due to the consistency of the TOD, it also seems sensible that more coherent bundles of tracks will run along the correct directions, hence establishing an increased amount of VC. An *a priori* assumption on the decrease of IC, however, is not *that* evident. As the amount of IC for the FOD-based tractography is already quite low, the margin for improvement is rather slim. Also, the ability of the TOD to guide tracks along greater lengths could have possibly led to an adverse effect: tracks that made a "wrong turn" before (when guided by the FOD), yet were unable to connect a pair of regions, simply ended up in the NC category; on the other hand, if they are now able to run across a longer distance (guided by the TOD), they risk ending up in the IC category. However, in practice, the TOD's consistency seems to overcome this risk by actually fixing enough of these "wrong turns"; the amount of IC still decreases clearly for the deterministic tractography, and stays almost the same (slight decrease) for the probabilistic tractography. Consequently, large increases of $\mathrm{VC}/(\mathrm{VC}+\mathrm{IC})$ are also observed. The filtered VC are visualized in Fig. 6.20 (the colors match those in Fig. 6.18). Even after only these valid tracks are filtered and observed on the more global level of whole bundles, the qualitative differences between FOD-based and TOD-based tractography are still apparent: the TOD-based results show *e.g.* more coherently delineated bundles, less bulging of bundle "6" into "7", etc. Finally note that, if we had not imposed the minimum track length of 24 mm, the rejected tracks would have ended up mostly in the NC category (they are too short to establish a valid connection; some might end up in the IC category though). As there are much more rejected tracks for the FOD-based experiments, this would – once again – render the differences between FOD-based and TOD-based tractography results even more contrasting.

6.4 Discussion

6.4.1 TODI: mapping the TOD in 5D

We introduced TODI as an extension of TDI in this work: TODI maps the *presence of tracks* in the 5D spatio-angular domain. When applied to a dense tractogram as obtained from any given fiber tractography procedure with associated parameters and constraints, the resulting TOD becomes representative of the full track distribution of which the tracking algorithm yields a finite sample.

Extending TDI towards the spatio-angular domain was also the aim of a previously proposed technique, called angular track imaging (ATI) (Pannek et al., 2012). At the core of the TODI and ATI mapping strategies, however,

there are two key differences. ATI represents a track in each voxel by considering only its single average tangent direction (*i.e.* it allows no representation of track curvature at the voxel level), and then maps this direction by an angular binning approach onto the closest sampling vector from a predetermined fixed set. TODI, on the other hand, considers all directional information of a track in each voxel by integrating over it, and then maps this directional information directly and unmodified to the TOD by using the mechanics of a KDE in the SH basis (Fig. 6.2). As we already argued in the introduction, we specifically avoided any particular choice of discrete angular bins due to an associated risk of bias, caused by typical differences in shape and size of such bins (Saff and Kuijlaars, 1997). Losing the ability to represent curvature of an individual track at the voxel level by employing only its single average tangent direction, might be a feasible option to reduce computation time at high spatial resolutions. Considering that the current limited resolution of DWI data causes voxels to be typically larger than $2 \times 2 \times 2\,\mathrm{mm}^3$, however, it is reasonable to expect curving tracks at the native resolution of the data (and even more so at sub-resolution). Depending on the application, it might not be advisable to ignore curvature in such a case.

The TOD as resulting from TODI is a true extension of the TDI map: when integrating the TOD over the angular domain, the TDI map is obtained. Furthermore, TODI "up to SH order 0" simply equates to TDI itself. In both TDI and TODI, each track delivers a *unit* contribution to each voxel. While this is a clear and simple definition, it also falls short in the sense that different tracks' intersections with a given voxel can have different lengths: not all tracks necessarily occupy the same partial "volume" within the voxel, yet contribute equally. This could be alleviated by weighting the contribution of a track to a voxel with the length of the intersecting part. For TODI, this simply results in canceling out the denominator "$\varepsilon_2 - \varepsilon_1$" in Eq. (6.3) and Eq. (6.4). We chose not to apply such weighting in the current work though, as our current intent was to maintain compatibility with the original definition and interpretation of TDI. It would however constitute a justified modification to both TDI and TODI.

We reasoned that TODI is, just like TDI, amenable to spatial super-resolution. As shown in Fig. 6.5F, this can aid in visual assessment of the structure of a complex track distribution. The description of geometrical and structural properties of DWI data or tractograms has also been the focus of other works, investigating *e.g.* crossing angles (Pasternak et al., 2012), track dispersion (Savadjiev et al., 2012), tract coherence (Vos et al., 2012), etc. Similar structural measures can be calculated based on the TOD in each voxel. As the TOD efficiently captures the main features of even the most complex probabilistic track distributions, we envision that it might additionaly prove to be a more convenient

(or supporting) representation for tackling difficult problems involving such distributions, *e.g.* track set registration or segmentation. Describing tractogram properties or solving complex problems involving tractograms can also readily be done in a multi-scale sense by obtaining the TOD at different spatial resolutions: due to the integration of angular information along the tracks in each voxel, an accurate summary of such properties remains encoded in the TOD. As such, there is also value in spatial sub-resolution instances of the TOD. The bins used in the spatial domain don't even have to take on the form of voxels: segmentations of certain structures could also be used to investigate the angular distribution of tracks within. In this work, we provided an example of the most extreme case of sub-resolution, *i.e.* encapsulating a complete tractogram in a single voxel. This yields a global representation of the angular distribution of the tracks. Note that the mapping strategy of ATI lacks the accuracy to be applied at this scale: each complete track would only be represented by a single average tangent direction. TODI nicely integrates along the full track lengths to map their angular profiles. We should, however, warn the reader that the presented result in Fig. 6.6 constitutes nothing more than an example of this possibility. Even though the outcome proved to be quite intriguing and well-structured, it was only based on a single subject. Furthermore, as we did not weight the contribution of each track with its length (as suggested above), the shorter tracks in the distribution might have been overrepresented (on the other hand, this might not have been too much of a problem, since it was a short-tracks distribution; *i.e.* the track lengths were independent of the tract lengths).

Finally, as the TOD does exactly quantify any tractogram itself, it could even be a useful tool to quantify and gain insight into the (sometimes subtle) differences between the track distributions resulting from different fiber tracking strategies or due to different parameter choices: it could as such also play an interesting role in the development and validation of new tractography algorithms.

Summary

- TODI maps the spatio-angular *presence of tracks*.
- TODI integrates all directional information along (curving) tracks within a voxel and maps it directly to the TOD by using a KDE mechanic.
- TODI is a true extension of TDI; in both methods, each track delivers a *unit* contribution to each voxel and integrating over the angular domain of the TOD yields the TDI map.

- TODI can provide an accurate description of any tractogram at any desired resolution, which may assist in solving other problems involving (complex) tractograms (*e.g.* registration or segmentation).
- TODI can also provide new insights into specific tractography algorithms and assist in their development and validation.

6.4.2 On meaning and (over)interpretation

Motivated by the spatio-angular super-resolution properties of the TOD and the resulting (often impressive) maps, it may be tempting to employ TODI for direct quantification of some aspects of the DWI data. Even though the TOD directly and accurately quantifies many aspects of any *tractogram*, it should however be kept in mind that most classical fiber tracking algorithms do *not* yield track densities that are related to any underlying biological property of the tissue under investigation. We have in fact shown recently that many features of a regular super-resolution TDI map depend on noise in the original DWI data, rather than true anatomical structure (Dhollander et al., 2012). Another study that included an investigation of the within-subject variability of TDI also concluded that TDI (and by extension, the amplitudes of the TOD resulting from TODI) is not suited for direct quantification of the acquired data (Willats et al., 2012). The problem is not with TDI (or TODI) itself, but rather with the preceding fiber tractography. Indeed, many different tracking strategies lead to very different results; a fact that is of uttermost importance when it comes down to interpretation of the results, especially impacting the domain of connectomics where interpretation is a key issue (Bastiani et al., 2012). As the number of available techniques is rapidly growing, leading scientists in the general domain of DWI are still emphasizing the issue of overinterpretation of track counts (Jones et al., 2013).

However, promising new methods are being proposed to improve the biological plausibility of tractograms by incorporating anatomically-constrained tractography (ACT) priors in the tracking process (Smith et al., 2012) and applying spherical-deconvolution informed filtering of tractograms (SIFT) for an optimal match between track density and AFD (Smith et al., 2013b,a). The TOD resulting from such tractograms could then become a more quantitative representation that combines the biological meaningfulness of AFD with the increased spatio-angular consistency provided by the fiber tractography.

In this work, we applied TODI to probabilistic *short-tracks* distributions. This

choice mitigates the dependence of the track densities on the underlying tract lengths, resulting in more uniform track densities (Fig. 6.3). Due to typical constraints such as the minimum track length and maximum amount of curvature, a track can only be present at a certain location in the distribution if a minimum amount of consistency with the neighborhood along the track can be established over a minimum distance. Continuity constraints have also been used in other recent works, such as Reisert and Kiselev (2011) and Reisert et al. (2012), for the reconstruction of a more regionally consistent FOD. The original CSD framework has also recently been extended to incorporate neighborhood information during the estimation of the FOD (Tournier et al., 2013b). The symmetric FOD as resulting from CSD, however, has a very different meaning as compared to the TOD. It encodes the AFD, *i.e.* the intra-axonal partial volume fraction of axons along different directions within a voxel (Raffelt et al., 2012b). The asymmetric FOD introduced in Reisert et al. (2012) bears yet another meaning: it is related to the amount of white matter fibers that smoothly turn into certain directions. In order to obtain this information, the latter work takes into account curvature at the voxel scale, but also assumes it to be constant. Our TODI method itself does not assume anything, yet the final resulting TOD takes on the assumptions of the fiber tractography method that generated the tractogram. This allows for the incorporation of several intuitive constraints (*e.g.* a minimum track length) at the level of the definition of the fiber tracking algorithm and its parameters. Their often complex combined impact on the resulting track distribution as a whole is automatically captured in the TOD.

We explained how the amplitude of a probabilistic short-tracks TOD can be understood as a measure of *track-like local support* (TLS). While it may be tempting to regard short-tracks TODI as a regularization mechanism (in the sense of "smoothing") based on the FOD, it is much closer in spirit with the technique of threshold-free cluster enhancement (TFCE) (Smith and Nichols, 2009). TFCE takes a raw statistic image and yields a map that represents the amount of cluster-like local spatial support in each voxel. Through its mechanism, voxels situated within cluster-like structures obtain an amount of support from surrounding voxels that are part of the same cluster-like structure (at multiple scales). In our setting of short-tracks TODI, orientations within voxels get their support from the tracks that run along them. Since these tracks were typically started from a seed point, one could reason that this seed point provided the support. Due to the maximum bidirectional tracking distance, support can only be received from seeds along the tract that are not too far away. As reasoned before, not enforcing a maximum distance would make the amount of support dependent on the tract length. Manipulating other parameters, such as the minimum track length and the maximum amount of curvature, allows for extra control over the definition of the TLS amplitude of the TOD. Reasoning about TLS as arising from the seed points of short tracks in the surroundings

allows for increased insight into its behavior. Most voxels in a single fiber region thus typically receive their support from both sides along the tract. Voxels within the maximum (bidirectional) tracking distance from a tract ending, will gradually receive less support from one direction, up to the point where voxels at the tract ending only receive half the support. Voxels in crossing regions receive much more support, because seeds located in several tracts may contribute; but the total amount of support is nicely divided between the lobes associated to these different tracts. Do note that all these insights rely on the fact that seeds are distributed uniformly over the brain. A different seeding strategy will of course also alter the interpretation of the resulting TLS measure. Finally, it is very important to clarify that, while the concept of TLS shares some common *spirit* with TFCE on a *highly abstract* level, it is anything but an *equivalent* of TFCE in the 5D domain. The idea of TFCE is to take a statistic and boost its original amplitude for greater belief in extended areas (*i.e.* clusters), yielding a new statistic. The TLS amplitude of the short-tracks TOD in this work, on the other hand, is by no means a modulation of the original AFD amplitude of the FOD: it merely indicates track-like structure as a geometrical phenomenon. For a true adaption of TFCE to be applied to a statistic as resulting from *e.g.* groupwise comparison of AFD, we refer the interested reader to Raffelt et al. (2013). In the latter work, the concept of spatial connectedness that is required as a part the original TFCE statistic's definition of a cluster, is redefined by a concept of structural connectedness which is actually based on the number of tracks. The overall common spirit between these techniques and TLS only lies in the fact that (part of) the support arises from the neighborhood, which is itself defined in a very specific way so as to match with a certain kind of geometrical structure (*e.g.* cluster-like or track-like structure) that is of interest.

The TLS interpretation of the short-tracks TOD amplitude allows to intuitively understand the success of the TOD-based tracking results. The TOD-based tractography we presented is a step-wise method that only takes into account the TOD at the current position to decide upon its next step, but on the other hand this TOD has information on track-like structure in the surroundings: it can guide a track along directions that are more likely to correspond to *continuous structure over a longer distance*. In yet another sense of interpretation, the initial massive probabilistic short-tracks tractography that leads to the TOD can be seen as an "exploratory" preprocessing step; as if the short tracks perform the field work and the final tracks profit! This essential information provided by the complex massive population of short tracks is captured and summarized within the TOD by the mechanism of TODI. Fiber tractography based on the TOD doesn't only lead to more robust results; it also renders the final tracking more easy, *i.e.* fewer tracks fail the constraints and more tracks travel a longer distance (along a plausible path). In all our tractography experiments, we required a certain total amount of tracks to be generated: the algorithm kept

on generating tracks until the required number of tracks was accepted. If more tracks are being rejected through the process, it simply takes longer to reach the required amount of accepted tracks. Consistently much fewer tracks were rejected using TOD-based tractography as compared to FOD-based tractography. The improvement in final quality of the result was most pronounced in data of lower quality, where the FODs presented more spurious peaks and their main lobes suffered from greater deviations. The TOD was overall able to overcome many small to large imperfections in the FODs resulting from data of different quality. This is in some sense similar to certain abilities of global tractography strategies (Fillard et al., 2009; Reisert et al., 2011) that can also overcome imperfections by taking into account a more global view of the tracks. In such methods, a track segment receiving enough support from surrounding structure (segments) that it is able to connect with in a continuous manner (*i.e.* by optimizing an internal bending/connectivity energy), can compensate for a reduced correspondence with the underlying data in cases where the latter is imperfect or lacks otherwise in providing the necessary information content. TOD-based tracking also shares a conceptual similarity with the 2^{nd} order integration over FODs (iFOD2) (Tournier et al., 2010) strategy. iFOD2 is based on the FOD, but takes the joint probability along each candidate path into account at each step. Such a mechanic looks one step ahead along the track to ensure better continuity with the FOD at the following step. This results in tracks that veer less off course and are able to follow tighter turns more robustly. The iFOD2 method could actually also be applied directly to TODs – *i.e.* 2^{nd} order integration over TODs (iTOD2) – for even more robust results. A further improvement is possible by already calculating the TOD itself based on a short-tracks probabilistic tractogram resulting from the iFOD2 method. Yet another similar method was recently proposed by Rowe et al. (2013). This method essentially takes the iFOD2 method further by probing along a longer distance ahead of the tracking through use of a particle filter mechanism. The advantage of our method is that the TOD only has to be calculated once as a preprocessing step. Tracking (targeted or whole-brain) can then proceed using existing simple (or more complex) algorithms. Different fiber tractography methods can indeed lead to a TOD with a certain (TLS) interpretation, that can in turn serve as the source for another tracking strategy. For instance, we could envision a TOD obtained from global tractography, or guiding it, or even the combination of both. The method proposed by Rowe et al. (2013) could profit of the increased consistency of the short-tracks TOD, but add unique value by resolving fanning polarity. Combinations with any of the other aforementioned tractography strategies are also perfectly possible. We should, however, stress the need for careful interpretation of such advanced constructions. We found that our simple but elegant approach of using probabilistic short-tracks to construct the TOD, and allowing the TOD to guide the same (probabilistic) tracking algorithm again

to obtain the final (targeted) tractograms, already yielded greatly robustified results; yet through a mechanism that is still reasonably comprehensible as a whole.

Summary

- The TOD directly quantifies any *tractogram*. However, keep in mind that many classical tractography algorithms currently do *not* yield biologically quantitative track densities. Beware overinterpretation of track counts.
- Recently proposed new methods may overcome this limitation and yield quantitative tractograms. The resulting TOD may combine the interpretation of AFD with the increased coherence of fiber tractography.
- The presented (short-tracks) TOD's coherence is the result of inherent continuity constraints of the preceding tractography. Continuity constraints have also been successfully used in other recent works. The TOD's definition is intuitively manipulated at the level of the fiber tracking algorithm.
- The amplitude of a probabilistic short-tracks TOD is a measure of *track-like local support* (TLS). A voxel receives (angular) support from seeds in the neighborhood, whose tracks are able to reach the voxel. The maximum tracking distance renders the amount of support independent of tract length.
- The TLS amplitude of the TOD can guide tracks along directions that are more likely to correspond to *continuous structure over a longer distance*. This explains the overall success and robustness of TOD-based tractography. The TOD only has to be calculated once as a preprocessing step.

6.4.3 Pipelines, choices and parameters

As is evident from the previous explanations, the TOD can play a role in different parts of a complex processing pipeline; due to its "compatibility" with many other processing steps. We hereby provide a sample scenario of such a state-of-the-art pipeline, as applied to *e.g.* a single dataset. Let's say we first obtain the FOD through application of CSD. Should the data be of challenging

quality, we can readily employ the latest improvements to the CSD framework to already make use of neighborhood information at this stage (Tournier et al., 2013b). We then generate a massive probabilistic short-tracks distribution using iFOD2 tracking. Next we apply TODI to this distribution to obtain the TOD, yielding information on TLS. This TOD can now be used to guide regular whole-brain fiber tracking (*e.g.* by iTOD2) that incorporates ACT priors (Smith et al., 2012). Due to the TLS interpretation of the TOD's amplitude, the tractography could more easily generate a higher amount of tracks that fulfill the strict anatomical constraints (*e.g.* connecting pairs of GM regions). Furthermore, the obtained tractogram can be subjected to SIFT (Smith et al., 2013b), using the original FOD for a better match between the resulting densities and the AFD. This should result in an improved biological interpretability of the tractogram (Smith et al., 2013a). Again the previous increased ease of generating a higher number of more robust tracks by TOD-based tractography is useful, since the SIFT process reduces the total amount of tracks (so it should start off from a very high number of tracks if a reasonably high amount is still to be retained in the end). The currently obtained result is a tractogram that features robust and continuous tracks between different GM regions, while presenting a distribution of densities that closely matches the AFD: a perfect candidate for a *meaningful* whole-brain connectome. Also, due to the key role of the TOD-based tractography in the process, the tractogram should consist of a very well behaved collection of coherent bundles with an absence of NC (thanks to the ACT priors), a minimal appearance of IC, and an overwhelming amount of VC. Furthermore, these properties should render it more suited for other post-processing operations such as automated tractogram segmentation (Smith et al., 2011). Finally, we can apply TODI to the tractogram again: the resulting TOD's amplitude can now be interpreted as a biologically meaningful quantitative measure (similar to AFD, but incorporating the different sources of prior information that were applied along the pipeline). Given that a sufficiently high amount of tracks were still retained after the SIFT process, it might even become *truly* justified to generate this TOD at super-resolution!

Experimenting with such advanced setups is far beyond the intention of our current work though: our aim was to focus on the role of the TOD itself, and additionally compare it as directly as possible to the FOD. Therefore, the targeted tractography experiments performed on both FODs and TODs were intentionally kept simple: for each experiment, only a single seed region was specified. Even quite straightforward adaptions could have easily yielded better results. For instance, we could have seeded whole-brain followed by selecting tracks through the region rather than only seeding from the region itself. Furthermore, if our interest had been to obtain good segmentations of known anatomical structures, more inclusion regions would certainly constrain the tractography much better; some useful setups of regions are provided in

e.g. Tournier et al. (2012). In the case of the *Fiber Cup* phantom, additional preprocessing steps and other choices along the pipeline would certainly also have the potential to improve the final results even more. Such further additions, however, would only distract from the direct comparison between the FOD and TOD we targeted in this work. To allow for the most fair comparison possible, we performed both FOD- and TOD-based tractography by always using the exact same algorithms and parameters for both. Due to a difference in orders of magnitude between the amplitudes of the FOD and the TOD, we had to apply different thresholds though. A realistic FOD threshold for a wide range of data and applications is typically on the order of 0.1–0.2. The TOD threshold depends on the average density of the underlying tractogram. In this work, we took on a rather pragmatic approach to determine appropriate TOD thresholds for each experiment: we compared the average amplitudes of both FOD and TOD in several regions to get an idea on the relative difference in order of magnitude between both. This allowed us to translate the realistic FOD threshold range of 0.1–0.2 to comparable magnitudes for the TOD. While this is certainly not the most optimal way to determine these values, this rather ad hoc approach already automatically yielded greatly improved results. We did *not* fine-tune the TOD thresholds any further beyond this first "guess", as we felt this would only bias the comparison in favor of the TOD-based tractography and thus compromise its fairness. The exact absolute threshold values used in our experiments should however not be directly applied to other cases. Because a good threshold depends on the average density of the underlying tractogram, factors such as the number of tracks, average track length, brain volume and others come into play. A way to mitigate dependency on these factors would be to normalize the TOD's magnitude by the average amplitude of the TOD in a reference region; similar to how the FOD is inherently normalized through the CSD process if the kernel is estimated from a reference region (*i.e.* all single fiber voxels) of the data itself. For the TOD, this normalization should be performed by simple scalar division by this reference magnitude rather than spherical deconvolution though. The resulting globally normalized TOD then comes into a sort of standard range, for which good reference thresholds can be established.

The TOD was also taken to the conceptual "next level" by using TOD-based tractography to generate a new probabilistic short-tracks tractogram that in turn yields a "level 2" TOD and accompanying "level 2" DEC TDI map. The relative improvements obtained by this second application of TODI were much smaller than those obtained from its initial application. Such a trend was also confirmed by the level 2 TOD-based tractography experiments. This makes sense, as both short-tracks distributions along the pipeline were always obtained using the same key tractography parameters. One could reason that the first pass of TODI already exploited most of the room for improvement and yields

a good measure of TLS. The difference between level 1 and 2 did become quite apparent in the super-resolution DEC TDI maps though. This is not surprising, as they are simply visualizing (up to great detail) a whole-brain short-tracks tractogram as respectively originating from FOD-based and (level 1) TOD-based tractography. The latter one provided a more regularized super-resolution short-tracks DEC TDI map. The differences between level 1 and 2 were most pronounced in data of lower quality; *i.e.* in more *challenging* conditions. Another interesting question for future research would be whether TOD-based tractography can provide additional value under other "challenging" conditions, such as the presence of different pathologies. In this context, but also in general, we should warn that iterating TODI many times in the hope to achieve some optimal result, might not be a good idea at all. While tractography is good at segmenting the continuous flow and coherent boundaries of bundles, it performs less well at determining the exact extent of these bundles (both width and length, *i.e.* termination points). This is due to the fact that choosing some FOD/TOD threshold is still part of the process: a slightly lower threshold will always result in slightly wider/longer bundles, and vice versa. Iterating TODI can thus result in bundle sizes that gradually deviate from their original size. In case of certain small pathologies the size of a few voxels, the bundles might start to "overgrow" and by consequence hide such pathologies. Therefore, we actually advise to stick with level 1 TOD-based tractography for most purposes. For the typical quality of most datasets, this will be sufficient and already improve upon FOD-based tractography by a large extent. The very small gains from level 2 TOD-based tractography might not be worth the extra computation time in such a case anyway. If level 2 or higher TOD-based tractography is used though, the results should be carefully checked and kept under control if necessary (*e.g.* by ACT priors).

We always applied TOD-based tractography by employing the TOD at the *native* resolution of the data. At this resolution, the short-tracks TOD as mapped by TODI is able to overcome small "defects" of the data at the scale of a single voxel. This proved to be optimal for noisy datasets where each voxel's FOD is independently affected by the noise. If a single voxel's FOD is affected up to the extent that it would severely interrupt the normal flow of tracks, yet the FOD in the surrounding voxels is sufficiently intact along the path of the tract, the voxel will still be able to get a good amount of TLS from the neighborhood along the correct directions. In our experience, TOD-based tractography guided by a super-resolution TOD actually performs less well (as it partially lacks the aforementioned property); but still better than FOD-based tractography. Additionally taking into account our previous warnings against overinterpretation of super-resolution TDI maps and TODs, we advise to stick with the (level 1) short-tracks TOD at the native resolution of the data for most practical purposes. The fact that the TOD's resolution can be freely chosen, can

be put to good use for datasets featuring anisotropic voxel sizes though: these can be made isotropic when TODI is applied along the pipeline. For instance, suppose a dataset with a slightly anisotropic resolution of $2.2 \times 1.9 \times 1.9\,\mathrm{mm}^3$: it would then be sensible to map the TOD at an isotropic resolution of *e.g.* $2 \times 2 \times 2\,\mathrm{mm}^3$. Staying within the same order of magnitude should still allow for the extra advantages of TOD-based tractography as discussed in this paragraph. Another advantage of not using super-resolution is the lower amount of short-tracks that are needed. In this work, we always used 80 million tracks for all our short-tracks distributions. This number was chosen to even allow for a very robust visualization of the super-resolution TOD of the *in vivo* HARDI dataset; yet was also kept for all the other experiments. However, for the purpose of using the native resolution TOD for TOD-based tractography, a lot fewer tracks are needed in practice, as is evident from the (absurdly) high TOD thresholds used in this work. In our experience, even fewer than 10 million short tracks are already sufficient (when using the same minimum and maximum track lengths and spatial resolution) to guarantee optimal robustness. For the phantoms, we could even do with much fewer tracks.

Finally, while the focus of this work is on TODI, the presented mechanism is fully compatible with that of TWI. Both techniques are direct extensions of TDI and their combination can result in a generalized framework of TOWI. To illustrate the compatibility of TODI with TWI and by consequence yield a proof of concept of TOWI, we provided the necessary basic formalisms and a short discussion in Appendix B.

Summary

- TODI can play different useful roles in complex processing pipelines. It integrates very well with existing techniques.
- We compared FOD-based and TOD-based tractography under equal conditions and using the exact same parameters. Since the TOD's amplitude is of a different order of magnitude though, we had to apply a different threshold. It was determined by "educated guessing", but better strategies can be thought of.
- When using the same parameters for FOD-based and subsequent level 1 TOD-based tractography, the extra gains for the resulting level 2 TOD at native resolution are relatively small compared to those already obtained at level 1. Special care should also be taken when iterating TODI multiple times in a row.

- Employing the (level 1) TOD at the native resolution of the data for TOD-based tractography, is able to overcome small "defects" of the data at the scale of a single voxel. TODI's freedom of choice for the resolution allows to work at an isotropic resolution, even if the original data features an anisotropic one.
- We advise to use the level 1 TOD at the native resolution of the data (or the closest isotropic resolution) for most TOD-based tractography purposes.

6.5 Conclusion

We introduced a method named track orientation density imaging (TODI), which aims to reconstruct the track orientation distribution (TOD) of any given tractogram. The method is formulated as a generalization of TDI, which only maps the 3D spatial distribution of tracks. The TOD as obtained from TODI, on the other hand, is a full 5D spatio-angular representation of all tracks. Just like the TDI map, the TOD is amenable to super-resolution, as well in the spatial as in the angular domain, but also to sub-resolution. As such, it allows for detailed mapping of a large population of tracks at multiple scales. In general, many interesting geometrical and structural properties of a tractogram are captured in a TOD, as it exactly quantifies the tractogram (but not necessarily any underlying biological properties of the data – this depends entirely on the tractography method).

We found that the amplitude of a probabilistic short-tracks TOD represents a measure of *track-like local support* (TLS). This motivated us to introduce the concept of TOD-based tractography. Such a strategy allows for greatly robustified tractography outcomes, especially in challenging conditions where the quality of the original data is limited: the short-tracks TOD can guide a track along directions that are more likely to correspond to *continuous structure over a longer distance*, *i.e. track-like* structure!

In conclusion, the TOD is a powerful descriptor of complex track distributions and allows such a distribution to be used for widely varying purposes: it is a versatile tool that can be cast in many different roles and scenarios in the expanding domain of fiber tractography based methods and their applications. We look eagerly forward to novel uses and specific interpretations of TODs based on the currently available as well as future tractography strategies.

Appendix A: Apodized delta functions

In Raffelt et al. (2012a), apodized delta functions were proposed as an alternative to the SH delta functions. By sacrificing some angular resolution (*i.e.* obtaining a wider main lobe) it is possible to drastically reduce the relative amplitude of side lobes, thus obtaining a better approximation of a spherical kernel with limited local support. An iterative strategy was also presented to calculate the SH coefficients of these apodized delta functions. We found, however, that the proposed algorithm in combination with the parameters provided in Raffelt et al. (2012a) does not always (*i.e.* for different maximum orders) robustly converge: for several maximum orders, it oscillates or doesn't (sufficiently) reduce the positive side lobes. In order to improve the robustness in generating apodized delta functions for a wider range of maximum orders, we changed a few details and some parameters. We hereby present the algorithm as adapted from Raffelt et al. (2012a).

The problem is to find the coefficients $\widetilde{e}_j$ that represent an apodized delta function $\widetilde{\delta}_n(\theta)$:

$$\widetilde{\delta}_n(\theta) = \sum_{j=0}^{n/2} \widetilde{e}_j Y_{2j}^0(\theta) \tag{6.9}$$

where we only need $n/2+1$ coefficients to go with the even order $Y_\ell^0(\theta)$, since $\widetilde{\delta}_n(\theta)$ only varies in function of θ. Note that a different indexing j, as compared to Eq. (6.1), is used here: $j = \ell/2$. The algorithm now proceeds as follows.

To calculate the even Y_ℓ^0 SH coefficient vector $\widetilde{\mathbf{e}}_n$ of $\widetilde{\delta}_n(\theta)$:

1. Generate a large set of uniformly distributed directions (*e.g.* by electrostatic repulsion, Jones et al. (1999a)), containing 1 direction oriented exactly along $\theta = 0$. We used 3000 directions.

2. Setup vector $\mathbf{d}$: $d_i = 1$ for the single $\theta = 0$ direction, $d_i = 0$ for all other directions.

3. Construct matrix $\mathbf{H}$ containing amplitudes of the even Y_ℓ^0 along all directions, for $0 \leq \ell \leq n$.

4. Initialize $\widetilde{\mathbf{e}}_n$ with the even Y_ℓ^0 SH coefficients of $\delta_n(\theta)$, by using Eq. (6.2). Then iterate until convergence:

   (a) Calculate amplitude vector $\mathbf{a}$ from coefficient vector $\widetilde{\mathbf{e}}_n$:

$$\mathbf{a} = \mathbf{H}\widetilde{\mathbf{e}}_n \tag{6.10}$$

(b) Construct diagonal regularization matrix $\mathbf{L}$:

$$\mathbf{L}_{i,i} = \begin{cases} 1 & \text{if } a_i < f \cdot \bar{a} \\ 0 & \text{otherwise} \end{cases} \tag{6.11}$$

where $\bar{a}$ is the average amplitude ($\bar{a} = \widetilde{e}_0/\sqrt{4\pi}$) and f is a certain fraction. We set $f = 0.1$.

(c) Obtain a new $\widetilde{\mathbf{e}}_n$ by solving regularized least-squares problem

$$(\mathbf{I} + \lambda\mathbf{L})\mathbf{H}\widetilde{\mathbf{e}}_n = \mathbf{d} \tag{6.12}$$

where $\mathbf{I}$ is the identity matrix and λ is a regularization weight. We chose $\lambda = 1000$.

5. Normalize $\widetilde{\mathbf{e}}_n$ by

$$\widetilde{\mathbf{e}}_n \leftarrow \frac{\widetilde{\mathbf{e}}_n}{\widetilde{e}_0 \cdot \sqrt{4\pi}} \tag{6.13}$$

so the final $\widetilde{\delta}_n(\theta)$ integrates to 1 over $\mathbb{S}^2$.

Apart from the *a priori* constraining of the algorithm to the even order $Y_\ell^0(\theta)$ basis functions, the final normalization step (Eq. (6.13)), and a couple different parameter choices, the most important difference is the definition of the regularization matrix (Eq. (6.11)). In Raffelt et al. (2012a), the regularization weight only applies for $a_i < 0$, *i.e.* the negative lobes that are still present at that particular iteration. In our version, it applies for all amplitudes below a certain fraction of the mean amplitude. This increases robustness to a large extent and in addition reduces the positive side lobes. The resulting apodized delta functions up to order 16 are shown in Fig. 6.2 (*first row*), and can be compared to the regular SH delta functions in Fig. 6.1 (*first row*). The algorithm always fully converged (*i.e.* $\mathbf{L}$, and thus also $\widetilde{\mathbf{e}}_n$, did not change any further), and successfully suppressed the full side lobe pattern. A detailed comparison between the amplitude profiles of the SH delta function and the apodized delta function for the $n = 8$ and $n = 16$ case, as well as all resulting coefficients for n up to 20 are provided in Fig. 6.21, Fig. 6.22 and Table 6.5.

To obtain an apodized delta function $\widetilde{\delta}_n^z(u)$ with its axis of symmetry along any arbitrary direction z, the following property can be used:

$$\widetilde{\delta}_n^z(u) = \widetilde{\delta}_n(\theta) * \delta_n^z(u) \tag{6.14}$$

where the spherical convolution can be computed by direct multiplication of the SH coefficients of $\delta_n^z(u)$ with the rotational harmonics (RH) coefficients of $\widetilde{\delta}_n(\theta)$ (Healy et al., 1998; Tournier et al., 2004, 2007). This property is key to obtaining the results in Eq. (6.7) and Eq. (6.8).

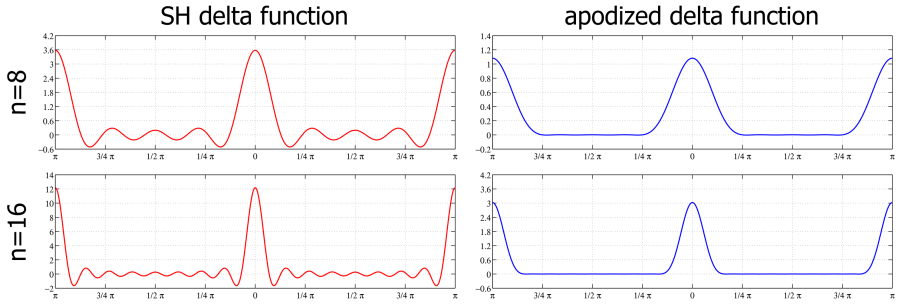

Figure 6.21: Comparison of the amplitude profiles of $\delta_n(\theta)$ and $\widetilde{\delta}_n(\theta)$, for $n = 8$ and $n = 16$. Note that the lobe of $\widetilde{\delta}_{16}(\theta)$ is comparable in shape (slighty sharper) to the main lobe of $\delta_8(\theta)$, through a combination of a higher maximum SH order (increases angular resolution) and apodization (decreases angular resolution).

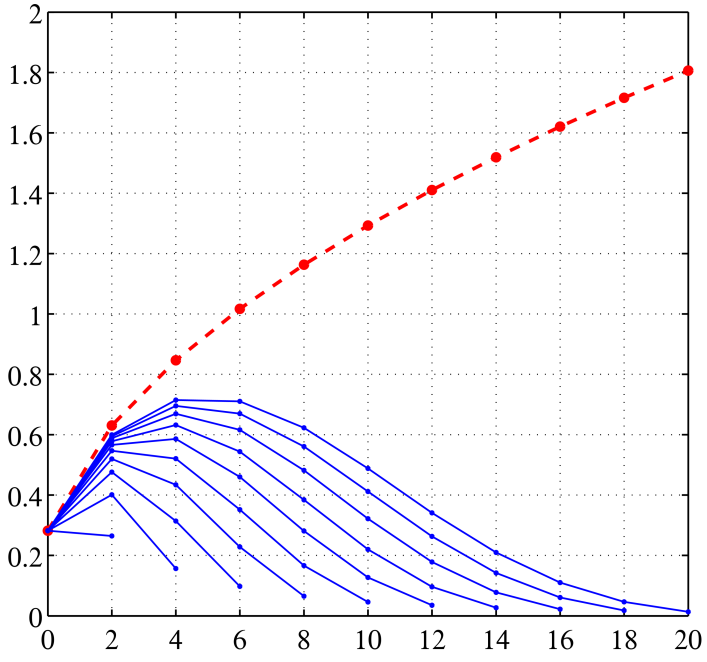

Figure 6.22: A plot of the even order Y_ℓ^0 coefficient vectors of δ_n (*red, dashed line*) and $\widetilde{\delta}_n$ (*blue, full lines*) for $n = 2 \ldots 20$. Note that, while the vector of δ_n can simply be truncated to the desired order, the elements of the vector of $\widetilde{\delta}_n$ depend on n. The actual values are provided in Table 6.5.

Table 6.5:

This table provides the values of the even order Y_ℓ^0 coefficient vectors of δ_n and $\widetilde{\delta}_n$ for $n = 2 \dots 20$.

ℓ	0	2	4	6	8	10	12	14	16	18	20
δ_n	0.2820947918	0.6307831305	0.8462843753	1.0171072363	1.1631066229	1.2927207365	1.4104739589	1.5191269449	1.6205112036	1.7159156294	1.8062879985
$\widetilde{\delta}_2$	0.2820947918	0.2645459007	–	–	–	–	–	–	–	–	–
$\widetilde{\delta}_4$	0.2820947918	0.4012319618	0.1564533344	–	–	–	–	–	–	–	–
$\widetilde{\delta}_6$	0.2820947918	0.4761803401	0.3141952797	0.0977918059	–	–	–	–	–	–	–
$\widetilde{\delta}_8$	0.2820947918	0.5196695259	0.4338197298	0.2282445903	0.0650534735	–	–	–	–	–	–
$\widetilde{\delta}_{10}$	0.2820947918	0.5470526360	0.5206293826	0.3516207639	0.1663283820	0.0458665394	–	–	–	–	–
$\widetilde{\delta}_{12}$	0.2820947918	0.5660506469	0.5862931541	0.4602398011	0.2811219643	0.1270249528	0.0352968569	–	–	–	–
$\widetilde{\delta}_{14}$	0.2820947918	0.5785407495	0.6320529393	0.5438245935	0.3841957934	0.2199412472	0.0960554650	0.0266596230	–	–	–
$\widetilde{\delta}_{16}$	0.2820947918	0.5882758938	0.6691135457	0.6159894424	0.4821624075	0.3216386667	0.1785835337	0.0775769693	0.0221270076	–	–
$\widetilde{\delta}_{18}$	0.2820947918	0.5950010962	0.6954831457	0.6698973538	0.5607479358	0.4118518254	0.2630266557	0.1420872623	0.0609982494	0.0175372490	–
$\widetilde{\delta}_{20}$	0.2820947918	0.5998215552	0.7147702826	0.7106419020	0.6230508551	0.4882538936	0.3413011721	0.2099196543	0.1102854613	0.0464721481	0.0132529913

Note that, while the vector of δ_n can simply be truncated to the desired order, the elements of the vector of $\widetilde{\delta}_n$ depend on n.

Appendix B: Track Orientation Weighted Imaging (TOWI)

Soon after the introduction of TDI (Calamante et al., 2010), a method to compute the average pathlength map (APM) (Pannek et al., 2011a) was proposed. Instead of counting all tracks intersecting a voxel, APM assigns the average length of those tracks to the voxel. The fact that both methods share the practice of obtaining a new contrast and super-resolution from a whole-brain tractogram, quickly led to the insight that they can be generalized to a common framework of super-resolution track-weighted imaging (TWI) (Calamante et al., 2012d). A similar extension was suggested by sampling diffusion indices along streamline trajectories (DIST) (Pannek et al., 2011b). More recently, other specific instances of the TWI framework were proposed that combine maps from functional magnetic resonance imaging (fMRI) analysis or positron emission tomography (PET) acquisitions with a tractogram from DWI in order to obtain new contrasts: track-weighted functional connectivity (TW-FC) (Calamante et al., 2012a, 2013) and track-weighted PET (TW-PET) (Calamante et al., 2012b).

In general, the TWI framework assigns a property P_{k_i} to each track k_i. This can be a global property of the track itself (*e.g.* its length), but it may also depend on the position along the track (*e.g.* the local curvature). It can even be a function of a property evaluated at all positions along the track (*e.g.* a sampling of another underlying, not necessarily diffusion related, map). In this case, the operation (*e.g.* sum, mean, maximum, ...) that combines these values along the track is referred to as the *track-wise* statistic and its output yields the value of P_{k_i}. Additionally, one can define a *grid-wise* statistic that combines the P_{k_i} values of all tracks k_i intersecting a particular voxel. The two most straightforward grid-wise statistics are the track-weighted total (TWT) and the track-weighted mean (TWM), respectively the sum and the mean. TDI is the specific case of $P_{k_i} = 1$ combined with the TWT. While TDI is as such a special case of TWI, we have shown it also is a special case of TODI; *i.e.* both TWI and TODI are generalizations of TDI. We will now introduce track orientation weighted imaging (TOWI), the integration of both the TWI and TODI extensions. The formalisms provided below should serve as a proof of concept, showcasing the compatibility of TWI and TODI.

All mechanisms and theory related to track-wise statistics remain unchanged and are extensively discussed in Calamante et al. (2012d). The grid-wise TWT statistic can be extended towards the definition of the track orientation weighted

total (TOWT), given by

$$\mathrm{TOWT}_n^P(r,u) = \sum_{i=1}^{K(r)} P_{k_i} \cdot C_n^{k_i}(r,u) \tag{6.15}$$

which is at the same time an extension of the TOD as defined in Eq. (6.5): the core TWI (TWT) and TODI mechanisms are combined as factors, mutually weighting each other. The case of $P_{k_i} = 1$ yields TODI, while $n = 0$ results in TWI (TWT). Setting both $P_{k_i} = 1$ and $n = 0$ simplifies the result to TDI. Just as we did for TODI, we can mitigate the Gibbs artifacts associated with the SH delta function by incorporating an apodized delta function into the definition. This results in

$$\widetilde{\mathrm{TOWT}}_n^P(r,u) = \widetilde{\delta}_n(\theta) * \mathrm{TOWT}_n^P(r,u) \tag{6.16}$$

through a simple derivation, analogously to Eq. (6.7) and Eq. (6.8). Again, we only need to compute the SH coefficients of $\widetilde{\delta}_n(\theta)$ once (see Appendix A).

Extending the grid-wise TWM statistic to an equivalent definition for the track orientation weighted mean (TOWM) introduces a new challenge. In Calamante et al. (2012d), the TWM is defined as the TWT divided by the total number of tracks $K(r)$ traversing the voxel. In order to obtain an expression for the TOWM, we can employ the TOWT; but we also need a generalization of $K(r)$ towards the angular domain. Because we didn't use an angular binning approach, there is no such integer track count available. However, just as $K(r)$ equals the TDI value, the angular version of $K(r)$ we're looking for is simply the TOD. Per consequence, a definition of the TOWM is obtained by

$$\mathrm{TOWM}_n^P(r,u) = \frac{\mathrm{TOWT}_n^P(r,u)}{\mathrm{TOD}_n(r,u)} \tag{6.17}$$

where it should be noted that the exact result can in general actually *not* be represented in the SH basis up to order n: the sets of SH coefficients of both $\mathrm{TOWT}_n^P(r,u)$ and $\mathrm{TOD}_n(r,u)$ need to be preserved. Because of the division by the TOD, the Gibbs artifacts are now a much bigger threat with respect to the quality of the result. It is therefore even more important to mitigate this by using the apodized delta function as the kernel of choice. The accompanying definition becomes

$$\widetilde{\mathrm{TOWM}}_n^P(r,u) = \frac{\widetilde{\mathrm{TOWT}}_n^P(r,u)}{\widetilde{\mathrm{TOD}}_n(r,u)} \tag{6.18}$$

which again needs two sets of SH coefficients to be exactly represented. This result is of a different form as compared to Eq. (6.8) and Eq. (6.16) because the

convolution with the apodized delta function appears in both the numerator and the denominator, *i.e.* the result does *not* equal $\widetilde{\delta}_n(\theta) * \mathrm{TOWM}_n^P(r,u)$. Just as for the TWM when $K(r) = 0$, the TOWM is not defined where the TOD equals 0. This can however also happen for certain regions of u in the angular domain, while $K(r) \neq 0$ in the voxel. Moreover, even though the apodized delta functions are *optimized* to have limited local support and constant zero amplitude over the rest of their domain, this is never *exactly* the case. In order to avoid numerical instabilities upon implementation, it is therefore advised to return 0 when evaluating the TOWM in (angular) regions where the TOD is smaller than a certain lower threshold. It might actually also be a good idea to do this for the original TWM when $K(r)$ has a very low value, to avoid mean values based on a very small number of (probably spurious) tracks.

In conclusion, we have provided the main formalisms necessary for the integration of TWI and TODI, resulting in a generalized framework of TOWI. Extending the TWT towards the TOWT proved to be rather trivial. The extension of the TWM towards the TOWM, however, poses some additional challenges that should be handled with appropriate care.

Chapter 7

Main contributions and research opportunities

In this chapter, a final brief overview is provided of the *main* contributions of this PhD thesis, as well as several opportunities for future research. For more elaborate and indepth discussion, the reader is referred to the discussions and conclusions provided with the previous individual chapters.

7.1 MSMT and PPVF

7.1.1 Main contributions

- A generic framework for multi-shell multi-tissue (MSMT) representations of DWI data consisting of (several) isotropic volume fractions (IVFs) and anisotropic volume fractions (AVFs).
- A specific instance of this framework, tailored for normal healthy (human) brains, that contains IVFs for CSF and GM, and an AVF for WM.
- A robust fitting strategy and regularization priors for the aforementioned MSMT representation, which encourage a non-negative fit and a sparse presence of tissue populations – without any assumptions on their individual appearances. This fitting procedure may also act as a fully automatic datadriven probabilistic tissue segmentation routine for DWI data, that doesn't require registration to an atlas prior – or even *any* other spatial assumptions.

- A novel DEC WM map based on the MSMT representation, which finally solves the main shortcomings of the classical DEC FA map from DTI.
- A novel preservation of principal volume fractions (PPVF) retransformation strategy, which profits from the properties of the MSMT representation to solve the retransformation problem trivially.
- New insights and a novel theory on transformational invariants of DWI data, allowing to identify biomarkers that are specific to non-geometrical aspects of fiber populations.
- A novel unbiased robust multi-subject multi-channel diffeomorphic registration algorithm; its application is illustrated for a group of healthy subjects using normalized volume fraction maps resulting from the MSMT representation as channels of information to guide the registration. The final outcome is combined with the PPVF strategy to obtain an average multi-shell DWI signal template in q-space.

7.1.2 Research opportunities

- Quantitative use of the MSMT fractions as direct (fiber population specific) biomarkers. The MSMT representation and PPVF strategy itself can act as the central mechanisms supporting the registration of groups of subjects to a common space in such studies.
- Development of novel pathology-specific biomarkers by studying the residuals from a regular MSMT fit, or even the addition of a specific volume fraction representing a certain pathology to the MSMT representation.
- Development of more "free-form" variants of existing models (*e.g.* NODDI) by encoding them as an MSMT representation.
- Development of more automated (iterative) approaches to obtain subject-specific tissue response functions.
- Exploration of other strategies to efficiently and reliably fit an MSMT representation.
- Taking PPVF "out of the box": greater flexibility while transforming DWI data by allowing individual tissue populations within a voxel to spatially "detach"; *i.e.* allowing individual bundles to translate independently. While this requires a new paradigm for DWI registration and transformation, the MSMT representation should provide a good basis for (reasoning on) such developments.

7.2 TODI and the TOD

7.2.1 Main contributions

- Track orientation density imaging (TODI): a natural generalization of TDI towards the 5D spatio-angular domain, that allows to map the complex track orientation distribution (TOD) of any tractogram.
- A specific formulation of TODI that allows for trivial use of any other kernel (besides the SH delta function) by a single convolution of the end result. This is specifically applied to obtain TODI with apodized delta functions, which does not suffer from Gibbs truncation artifacts.
- A generic theory that allows for spatial and angular, super-resolution and sub-resolution mapping of the TOD.
- Increased insight in the properties of short-tracks distributions, and the accompanying interpretation of the amplitude of short-tracks TODs as a measure of track-like local support (TLS). The TLS interpretation can be easily manipulated at the intuitive level of the tractography parameters. (*e.g.* minimum track length, amount of allowed curvature, ...)
- TOD-based tractography, which relies on the TLS amplitude to guide tracks along directions that are more likely to correspond to continuous structure over a longer distance, *i.e.* track-like structure!
- The novel concept of multi-level TDI and TODI.
- A generized framework of track orientation weighted imaging (TOWI), integrating both extensions of TDI, as offered by TODI and TWI.

7.2.2 Research opportunities

- Employing the TOD to assist in solving other complex tasks involving tractograms: track set registration, segmentation, ...
- Using the TOD to quantify differences between tractograms and investigate the effects caused by certain choices of tractography algorithms and parameters.
- Using the TOD to quantify other geometrical properties of tractograms. (*e.g.* number of fiber populations, crossing angle, amount of dispersion, ...)

- Integrating the TOD in state-of-the-art pipelines as a more reliable way of quickly generating large amounts of valid connections, or as a descriptor of biologically plausible tractograms.
- Performing TOD-based tractography under challenging conditions (*e.g.* low data quality, artifacts, ...), or in the presence of severe pathologies.
- Performing TOD-based tractography using TODs resulting from other tractography strategies, or employing the TOD to guide other tractography strategies, or even the combination of both.
- Defining several useful variants of TOWI to use in quantitative group studies.

7.3 Final conclusion

As I'm writing this final conclusion, I have the checklist of our doctoral school right next to me. The core sentence reads: *"The thesis must contain the necessary information allowing the examination committee to assess whether the doctoral researcher has developed into an independent, critical and creative researcher"*. As I'm expected to have developed these qualities, I hereby choose to use my academic freedom to make a final (personal) claim:

I think we are there. Anyone active in the field of DWI and who has been paying attention during the last 5 (or so) years, should know what I'm hinting at. The last years have been a constant battle to move truly beyond DTI. Don't get me wrong: DTI has been an important historical development that allowed us to be where we are now. But DTI is also lacking, misleading – and admit it – plain *wrong* in so many ways. However, through the combined efforts of many people working actively to provide alternative higher order models and accompanying techniques, I believe we have finally reached a point where we can do *everything* that DTI allowed us to do; but better, more correct in every single relevant aspect, and using clinically feasible acquisition setups. All the methods are there – from visualization to registration to segmentation (tractography) to quantification (just check the bibliography section) – to perform the whole pipeline from beginning to end, without the need to resort to *any* aspect of the deprecated DTI model. Now it's only a matter of actually doing so. I hope we can make this happen during the next 5 (or so) years; and that my work and efforts so far have contributed their own tiny bit towards this goal.

Acronyms

ACT	Anatomically-Constrained Tractography
ADC	Apparent Diffusion Coefficient
AFD	Apparent Fiber Density
APM	Average Pathlength Map
ATI	Angular Track Imaging
AVF	Anisotropic Volume Fraction
B0	Non diffusion weighted image ($b = 0$)
CB	Cingulum Bundle
CNR	Contrast-to-Noise Ratio
COD	Coefficient Of Dispersion
COV	Coefficient Of Variation
CSD	Constrained Spherical Deconvolution
CSF	Cerebrospinal Fluid
DEC	Directionally-Encoded Color
DIST	Diffusion Indices along Streamline Trajectories
DTI	Diffusion Tensor Imaging
DWI	Diffusion Weighted Imaging
EPI	Echo Planar Imaging
FA	Fractional Anisotropy

fMRI	Functional Magnetic Resonance Imaging
FOD	Fiber Orientation Distribution
GCC	Genu of the Corpus Callosum
GM	Gray Matter
HARDI	High Angular Resolution Diffusion Imaging
IC	Invalid Connections
iFOD2	2^{nd} order Integration over Fiber Orientation Distributions
iTOD2	2^{nd} order Integration over Track Orientation Distributions
IVF	Isotropic Volume Fraction
KDE	Kernel Density Estimate
LLS	Linear Least Squares
MD	Mean Diffusivity
MR	Magnetic Resonance
MRI	Magnetic Resonance Imaging
MSMT	Multi-Shell Multi-Tissue
NC	No Connections
ND	Neurite Density
NLS	Nonlinear Least Squares
NODDI	Neurite Orientation Dispersion and Density Imaging
OD	Orientation Dispersion
ODF	Orientation Density Function
PDF	Probability Density Function
PET	Positron Emission Tomography
PPD	Preservation of Principal Direction
PPVF	Preservation of Principal Volume Fractions
PSF	Point Spread Function

RESTORE	Robust Estimation of Tensors by Outlier Rejection
RGB	Red–Green–Blue
ROI	Region Of Interest
RH	Rotational Harmonics
SCP	Superior Cerebellar Peduncle
SH	Spherical Harmonics
SIFT	Spherical-deconvolution Informed Filtering of Tractograms
SNR	Signal-to-Noise Ratio
TDI	Track-Density Imaging
TFCE	Threshold-Free Cluster Enhancement
TLS	Track-like Local Support
TODI	Track Orientation Density Imaging
TOD	Track Orientation Distribution
TOWI	Track Orientation Weighted Imaging
TOWM	Track Orientation Weighted Mean
TOWT	Track Orientation Weighted Total
TRSE	Twice-Refocused Spin-Echo
TW-FC	Track-Weighted Functional Connectivity
TWI	Track-Weighted Imaging
TWM	Track-Weighted Mean
TW-PET	Track-Weighted Positron Emission Tomography
TWT	Track-Weighted Total
VC	Valid Connections
WLLS	Weighted Linear Least Squares
WM	White Matter

Bibliography

Alexander, D.C., Pierpaoli, C., Basser, B.J., Gee, J.C., 2001. Spatial transformations of diffusion tensor magnetic resonance images. IEEE Trans. Med. Imag. 20, 1131–1139.

Alexander, D.C., Seunarine, K.K., 2010. Mathematics of crossing fibers, in: Jones, D.K. (Ed.), Diffusion MRI: Theory, Methods, and Applications. Oxford Univ. Pr., pp. 451–464.

Andersson, J.L.R., Skare, S., Ashburner, J., 2003. How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. NeuroImage 20, 870–888.

Andersson, J.L.R., Xu, J., Yacoub, E., Auerbach, E., Moeller, S., Ugurbil, K., 2012. A comprehensive gaussian process framework for correcting distortions and movements in diffusion images, in: Proc. Intl. Soc. Mag. Reson. Med., p. 2426.

Barmpoutis, A., Vemuri, B.C., Forder, J.R., 2007. Registration of high angular resolution diffusion MRI images using 4th order tensors, in: Med. Image Comput. Comput. Assist. Interv., pp. 908–915.

Basser, P.J., 1995. Inferring microstructural features and the physiological state of tissues from diffusion-weighted images. NMR Biomed. 8, 333–344.

Basser, P.J., 1998. Fiber-tractography via diffusion tensor MRI (DT-MRI), in: Proc. Intl. Soc. Mag. Reson. Med., p. 1226.

Basser, P.J., Mattiello, J., Le Bihan, D., 1994a. Estimation of the effective self-diffusion tensor from the NMR spin echo. J. Magn. Reson. Ser. B 103, 247–254.

Basser, P.J., Mattiello, J., Le Bihan, D., 1994b. Mr diffusion tensor spectroscopy and imaging. Biophys. J. 66, 259–267.

Basser, P.J., Pajevic, S., Pierpaoli, C., Duda, J., Aldroubi, A., 2000. In vivo fiber tractography using DT-MRI data. Magn. Reson. Med. 44, 625–632.

Bastiani, M., Shah, N.J., Goebel, R., Roebroeck, A., 2012. Human cortical connectome reconstruction from diffusion weighted MRI: The effect of tractography algorithm. NeuroImage 62, 1732–1749.

Calamante, F., Masterton, R.A.J., Tournier, J.D., Smith, R.E., Willats, L., Raffelt, D., Connelly, A., 2012a. Super-resolution track-weighted functional connectivity (TW-FC): a tool for characterizing the structural-functional connections in the brain, in: Proc. Intl. Soc. Mag. Reson. Med., p. 139.

Calamante, F., Masterton, R.A.J., Tournier, J.D., Smith, R.E., Willats, L., Raffelt, D., Connelly, A., 2013. Track-weighted functional connectivity (TW-FC): A tool for characterizing the structural-functional connections in the brain. NeuroImage 70, 199–210.

Calamante, F., Son, Y.D., Tournier, J.D., Ryu, T.H., Oh, S.H., Connelly, A., Cho, Z.H., 2012b. Fusing PET and MRI data using super-resolution track-weighted imaging, in: Proc. Intl. Soc. Mag. Reson. Med., p. 1919.

Calamante, F., Tournier, J.D., Heidemann, R.M., Anwander, A., Jackson, G.D., Connelly, A., 2011. Track density imaging (TDI): Validation of super resolution property. NeuroImage 56, 1259–1266.

Calamante, F., Tournier, J.D., Jackson, G.D., Connelly, A., 2010. Track-density imaging (TDI): Super-resolution white matter imaging using whole-brain track-density mapping. NeuroImage 53, 1233–1243.

Calamante, F., Tournier, J.D., Kurniawan, N.D., Yang, Z., Gyengesi, E., Galloway, G.J., Reutens, D.C., Connelly, A., 2012c. Super-resolution track-density imaging studies of mouse brain: Comparison to histology. NeuroImage 59, 286–296.

Calamante, F., Tournier, J.D., Smith, R.E., Connelly, A., 2012d. A generalised framework for super-resolution track-weighted imaging. NeuroImage 59, 2494–2503.

Caruyer, E., Lenglet, C., Sapiro, G., Deriche, R., 2013. Design of multishell sampling schemes with uniform coverage in diffusion MRI. Magn. Reson. Med. 69, 1534–1540.

Chang, L.C., Jones, D.K., Pierpaoli, C., 2005. RESTORE: robust estimation of tensors by outlier rejection. Magn. Reson. Med. 53, 1088–1095.

Chang, L.C., Walker, L., Pierpaoli, C., 2012. Informed RESTORE: a method for robust estimation of diffusion tensor from low redundancy datasets in the presence of physiological noise artifacts. Magn. Reson. Med. 68, 1654–1663.

Close, T.G., Tournier, J.D., Calamante, F., Johnston, L.A., Mareels, I., Connelly, A., 2009. A software tool to generate simulated white matter structures for the assessment of fibre-tracking algorithms. NeuroImage 47, 1288–1300.

Conturo, T.E., Lori, N.F., Cull, T.S., Akbudak, E., Snyder, A.Z., Shimony, J.S., Mc Kinstry, R.C., Burton, H., Raichle, M.E., 1999. Tracking neuronal fiber pathways in the living human brain, in: Proc. Natl. Acad. Sci. USA, pp. 10422–10427.

Côté, M.A., Boré, A., Girard, G., Houde, J.C., Descoteaux, M., 2012. Tractometer: Online evaluation system for tractography, in: Med. Image Comput. Comput. Assist. Interv., pp. 698–705.

Côté, M.A., Girard, G., Boré, A., Garyfallidis, E., Houde, J.C., Descoteaux, M., 2013. Tractometer: Towards validation of tractography pipelines. Med. Image Anal. 17, 844–857.

Descoteaux, M., Angelino, E., Fitzgibbons, S., Deriche, R., 2006. Apparent diffusion coefficients from high angular resolution diffusion imaging: Estimation and applications. Magn. Reson. Med. 56, 395–410.

Dhollander, T., Emsell, L., Van Hecke, W., Maes, F., Sunaert, S., Suetens, P., 2012. Track-density imaging & noise: when super-resolution quality does not yield accuracy, in: Proc. Intl. Soc. Mag. Reson. Med., p. 1920.

Dhollander, T., Van Hecke, W., Maes, F., Sunaert, S., Suetens, P., 2010. Spatial transformations of high angular resolution diffusion imaging data in q-space, in: CDMRI Workshop at MICCAI, pp. 73–83.

Dhollander, T., Van Hecke, W., Maes, F., Sunaert, S., Suetens, P., 2011a. A hybrid diffusion imaging atlas in q-space, in: Proc. Intl. Soc. Mag. Reson. Med., p. 416.

Dhollander, T., Van Hecke, W., Maes, F., Sunaert, S., Suetens, P., 2011b. Methods for reorienting and retransforming diffusion weighted imaging data, in: Proc. Intl. Soc. Mag. Reson. Med., p. 4664.

Dhollander, T., Veraart, J., Van Hecke, W., Maes, F., Sunaert, S., Sijbers, J., Suetens, P., 2011c. Feasibility and advantages of diffusion weighted imaging atlas construction in q-space, in: Med. Image Comput. Comput. Assist. Interv., pp. 166–173.

Du, J., Goh, A., Qiu, A., 2012. Diffeomorphic metric mapping of high angular resolution diffusion imaging based on riemannian structure of orientation distribution functions. IEEE Trans. Med. Imag. 31, 1021–1033.

Farquharson, S., Tournier, J.D., Calamante, F., Fabinyi, G., Schneider-Kolsky, M., Jackson, G.D., Connelly, A., 2013. White matter fiber tractography: why we need to move beyond DTI. J. Neurosurg. 118, 1367–1377.

Fillard, P., Descoteaux, M., Goh, A., Gouttard, S., Jeurissen, B., Malcolm, J., Ramirez-Manzanares, A., Reisert, M., Sakaie, K., Tensaouti, F., Yo, T., Mangin, J.F., Poupon, C., 2011. Quantitative evaluation of 10 tractography algorithms on a realistic diffusion MR phantom. NeuroImage 56, 220–234.

Fillard, P., Poupon, C., Mangin, J.F., 2009. A novel global tractography algorithm based on an adaptive spin glass model, in: Med. Image Comput. Comput. Assist. Interv., pp. 927–934.

Fischl, B., 2012. FreeSurfer. NeuroImage 62, 774–781.

Glasser, M.F., Sotiropoulos, S.N., Wilson, J.A., Coalson, T.S., Fischl, B., Andersson, J.L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J.R., Van Essen, D.C., Jenkinson, M., for the WU-Minn HCP Consortium, 2013. The minimal preprocessing pipelines for the Human Connectome Project. NeuroImage 80, 105–124.

Hagmann, P., 2005. From diffusion MRI to brain connectomics. Ph.D. thesis. École Polytechnique Fédérale de Lausanne.

Hagmann, P., Cammoun, L., Gigandet, X., Gerhard, S., Grant, P.E., Wedeen, V.J., Meuli, R., Thiran, J.P., Honey, C.J., Sporns, O., 2010. MR connectomics: Principles and challenges. J. Neurosci. Meth. 194, 34–45.

Healy, D.M., Hendriks, H., Kim, P.T., 1998. Spherical deconvolution. J. Multivariate Anal. 67, 1–22.

Hong, X., Arlinghaus, L. R., Anderson, A.W., 2009. Spatial normalization of the fiber orientation distribution based on high angular resolution diffusion imaging. Magn. Reson. Med. 61, 1520–1527.

Jbabdi, S., Johansen-Berg, H., 2011. Tractography: Where do we go from here? Brain Connect. 1, 169–183.

Jbabdi, S., Sotiropoulos, S.N., Savio, A.M., Graña, M., Behrens, T.E.J., 2012. Model-based analysis of multishell diffusion MR data for tractography: How to get over fitting problems. Magn. Reson. Med. 68, 1846–1855.

Jenkinson, M., Beckmann, C.F., Behrens, T.E.J., Woolrich, M.W., Smith, S.M., 2012. FSL. NeuroImage 62, 782–790.

Jeurissen, B., Leemans, A., Jones, D.K., Tournier, J.D., Sijbers, J., 2011. Probabilistic fiber tracking using the residual bootstrap with constrained spherical deconvolution. Hum. Brain Mapp. 32, 461–479.

Jeurissen, B., Leemans, A., Tournier, J.D., Jones, D.K., Sijbers, J., 2013. Investigating the prevalence of complex fiber configurations in white matter tissue with diffusion magnetic resonance imaging. Hum. Brain Mapp. 34, 2747–2766.

Jones, D.K., 2010. Diffusion MRI: Theory, Methods, and Applications. Oxford Univ. Pr.

Jones, D.K., Horsfield, M., Simmons, A., 1999a. Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging. Magn. Reson. Med. 42, 515–525.

Jones, D.K., Knösche, T.R., Turner, R., 2013. White matter integrity, fiber count, and other fallacies: the do's and don'ts of diffusion MRI. NeuroImage 73, 239–254.

Jones, D.K., Simmons, A., Williams, S.C.R., Horsfield, M.A., 1999b. Non-invasive assessment of axonal fiber connectivity in the human brain via diffusion tensor MRI. Magn. Reson. Med. 42, 37–41.

Le Bihan, D., Breton, E., 1985. Imagerie de diffusion in vivo par résonance magnétique nucléaire. C. R. Acad. Sc. Paris T.301, Série II, 1109–1112.

Le Bihan, D., Breton, E., Lallemand, D., Grenier, P., Cabanis, E., Laval-Jeantet, M., 1986. MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders. Radiology 161, 401–407.

Lori, N.F., Cull, T.S., Akbudak, E., Snyder, A.Z., Shimony, J.S., Burton, H., Raichle, M.E., Conturo, T.E., 1999. Tracking neuronal fibers in the living human brain with diffusion MRI, in: Proc. Intl. Soc. Mag. Reson. Med., p. 324.

Mori, S., Crain, B.J., Chacko, V.P., Zijl, P.C.V., 1999. Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. Ann. Neurol. 45, 265–269.

Pajevic, S., Pierpaoli, C., 1999. Color schemes to represent the orientation of anisotropic tissues from diffusion tensor data: application to white matter fiber tract mapping in the human brain. Magn. Reson. Med. 42, 526–540.

Pannek, K., Mathias, J.L., Bigler, E.D., Brown, G., Taylor, J.D., Rose, S.E., 2011a. The average pathlength map: A diffusion MRI tractography-derived index for studying brain pathology. NeuroImage 55, 133–141.

Pannek, K., Mathias, J.L., Rose, S.E., 2011b. MRI diffusion indices sampled along streamline trajectories: Quantitative tractography mapping. Brain Connect. 1, 331–338.

Pannek, K., Raffelt, D., Salvado, S., Rose, S., 2012. Incorporating directional information in diffusion tractography derived maps: angular track imaging (ATI), in: Proc. Intl. Soc. Mag. Reson. Med., p. 1912.

Pasternak, O., Rathi, Y., Shenton, M.E., Westin, C.F., 2012. Estimation of the angle between crossing fibers as a novel structural quantity, in: Proc. Intl. Soc. Mag. Reson. Med., p. 1915.

Peter, D., Tape, C., Boschi, L., Woodhouse, J.H., 2007. Surface wave tomography: global membrane waves and adjoint methods. Geophys. J. Int. 171, 1098–1117.

Poupon, C., Clark, C.A., Frouin, V., Bloch, I., Bihan, D.L., Mangin, J.F., 1999. Tracking white matter fascicles with diffusion tensor imaging, in: Proc. Intl. Soc. Mag. Reson. Med., p. 325.

Poupon, C., Laribiere, L., Tournier, G., Bernard, J., Fournier, D., Fillard, P., Descoteaux, M., Mangin, J.F., 2010. A diffusion hardware phantom looking like a coronal brain slice, in: Proc. Intl. Soc. Mag. Reson. Med., p. 581.

Poupon, C., Rieul, B., Kezele, I., Perrin, M., Poupon, F., Mangin, J.F., 2008. New diffusion phantoms dedicated to the study and validation of high-angular-resolution diffusion imaging (HARDI) models. Magn. Reson. Med. 60, 1276–1283.

Raffelt, D., Smith, R.E., Tournier, J.D., Ridgway, G.R., Villemagne, V.L., Rowe, C.C., Salvado, O., Connelly, A., 2013. Tractographic threshold-free cluster enhancement: Whole-brain statistical analysis of diffusion MRI measures in the presence of crossing fibres, in: Proc. Intl. Soc. Mag. Reson. Med., p. 841.

Raffelt, D., Tournier, J.D., Crozier, S., Connelly, A., Salvado, O., 2012a. Reorientation of fiber orientation distributions using apodized point spread functions. Magn. Reson. Med. 67, 844–855.

Raffelt, D., Tournier, J.D., Fripp, J., Crozier, S., Connelly, A., Salvado, O., 2009. Non-linear spatial normalization of high angular resolution diffusion imaging data using fiber orientation distributions, in: DMFC Workshop at MICCAI, pp. –.

Raffelt, D., Tournier, J.D., Fripp, J., Crozier, S., Connelly, A., Salvado, O., 2011. Symmetric diffeomorphic registration of fibre orientation distributions. NeuroImage 56, 1171–1180.

Raffelt, D., Tournier, J.D., Rose, S., Ridgway, G.R., Henderson, R., Crozier, S., Salvado, O., Connelly, A., 2012b. Apparent fibre density: A novel measure for the analysis of diffusion-weighted magnetic resonance images. NeuroImage 59, 3976–3994.

Reese, T.G., Heid, O., Weisskoff, R.M., Wedeen, V.J., 2003. Reduction of eddy-current-induced distortion in diffusion MRI using a twice-refocused spin echo. Magn. Reson. Med. 49, 177–182.

Reisert, M., Kellner, E., Kiselev, V.G., 2012. About the geometry of asymmetric fiber orientation distributions. IEEE Trans. Med. Imag. 31, 1240–1249.

Reisert, M., Kiselev, V.G., 2011. Fiber continuity: An anisotropic prior for ODF estimation. IEEE Trans. Med. Imag. 30, 1274–1283.

Reisert, M., Mader, I., Anastasopoulos, C., Weigel, M., Schnell, S., Kiselev, V.G., 2011. Global fiber reconstruction becomes practical. NeuroImage 54, 955–962.

Rowe, M., Zhang, H., Oxtoby, N., Alexander, D.C., 2013. Beyond crossing fibers: Tractography exploiting sub-voxel fibre dispersion and neighbourhood structure, in: Inf. Process. Med. Imaging, pp. 402–413.

Saff, E.B., Kuijlaars, A.B.J., 1997. Distributing many points on a sphere. Math. Intel. 19, 5–11.

Savadjiev, P., Rathi, Y., Bouix, S., Verma, R., Westin, C.F., 2012. Multi-scale characterization of white matter tract geometry, in: Med. Image Comput. Comput. Assist. Interv., pp. 34–41.

Smith, R.E., Tournier, J.D., Calamante, F., Connelly, A., 2011. A novel paradigm for automated segmentation of very large whole-brain probabilistic tractography data sets, in: Proc. Intl. Soc. Mag. Reson. Med., p. 673.

Smith, R.E., Tournier, J.D., Calamante, F., Connelly, A., 2012. Anatomically-constrained tractography: Improved diffusion MRI streamlines tractography through effective use of anatomical information. NeuroImage 62, 1924–1938.

Smith, R.E., Tournier, J.D., Calamante, F., Connelly, A., 2013a. Evidence for the improved biological interpretability of white matter connectivity derived following tractogram filtering using SIFT, in: Proc. Intl. Soc. Mag. Reson. Med., p. 2135.

Smith, R.E., Tournier, J.D., Calamante, F., Connelly, A., 2013b. SIFT: Spherical-deconvolution informed filtering of tractograms. NeuroImage 67, 298–312.

Smith, S.M., Nichols, T.E., 2009. Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. NeuroImage 44, 83–98.

Sotiropoulos, S.N., Jbabdi, S., Xu, J., Andersson, J.L., Moeller, S., Auerbach, E.J., Glasser, M.F., Hernandez, M., Sapiro, G., Jenkinson, M., Feinberg, D.A., Yacoub, E., Lenglet, C., Van Essen, D.C., Ugurbil, K., Behrens, T.E.J., for the WU-Minn HCP Consortium, 2013. Advances in diffusion MRI acquisition and processing in the Human Connectome Project. NeuroImage 80, 125–143.

Sporns, O., Tononi, G., Kötter, R., 2005. The human connectome: A structural description of the human brain. PLoS Comput. Biol. 1, e42.

Stejskal, E.O., Tanner, J.E., 1965. Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. J. Chem. Phys. 42, 288–292.

Tao, X., Miller, J.V., 2006. A method for registering diffusion weighted magnetic resonance images, in: Med. Image Comput. Comput. Assist. Interv., pp. 594–602.

Tax, C.M.W., Jeurissen, B., Vos, S.B., Viergever, M.A., Leemans, A., 2014. Recursive calibration of the fiber response function for spherical deconvolution of diffusion MRI data. NeuroImage 86, 67–80.

Thirion, J.P., 1998. Image matching as a diffusion process: an analogy with Maxwell's demons. Med. Image Anal. 2, 243–260.

Tournier, J.D., 2010. The biophysics of crossing fibers, in: Jones, D.K. (Ed.), Diffusion MRI: Theory, Methods, and Applications. Oxford Univ. Pr., pp. 465–481.

Tournier, J.D., Calamante, F., Connelly, A., 2007. Robust determination of the fibre orientation distribution in diffusion MRI: Non-negativity constrained super-resolved spherical deconvolution. NeuroImage 35, 1459–1472.

Tournier, J.D., Calamante, F., Connelly, A., 2010. Improved probabilistic streamlines tractography by 2nd order integration over fibre orientation distributions, in: Proc. Intl. Soc. Mag. Reson. Med., p. 1670.

Tournier, J.D., Calamante, F., Connelly, A., 2012. MRtrix: Diffusion tractography in crossing fiber regions. Int. J. Imaging Syst. Technol. 22, 53–66.

Tournier, J.D., Calamante, F., Connelly, A., 2013a. Determination of the appropriate b value and number of gradient directions for high-angular-resolution diffusion-weighted imaging. NMR Biomed. 26, 1775–1786.

Tournier, J.D., Calamante, F., Connelly, A., 2013b. A robust spherical deconvolution method for the analysis of low SNR or low angular resolution diffusion data, in: Proc. Intl. Soc. Mag. Reson. Med., p. 772.

Tournier, J.D., Calamante, F., Gadian, D.G., Connelly, A., 2004. Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. NeuroImage 23, 1176–1185.

Tournier, J.D., Yeh, C.H., Calamante, F., Cho, K.H., Connelly, A., Lin, C.P., 2008. Resolving crossing fibres using constrained spherical deconvolution: Validation using diffusion-weighted imaging phantom data. NeuroImage 42, 617–625.

Tuch, D.S., Reese, T.G., Wiegell, M.R., Makris, N., Belliveau, J.W., Wedeen, V.J., 2002. High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. Magn. Reson. Med. 48, 577–582.

Ugurbil, K., Xu, J., Auerbach, E.J., Moeller, S., Vu, A.T., Duarte-Carvajalino, J.M., Lenglet, C., Wu, X., Schmitter, S., Van de Moortele, P.F., Strupp, J., Sapiro, G., De Martino, F., Wang, D., Harel, N., Garwood, M., Chen, L., Feinberg, D.A., Smith, S.M., Miller, K.L., Sotiropoulos, S.N., Jbabdi, S., Andersson, J.L.R., Behrens, T.E.J., Glasser, M.F., Van Essen, D.C., Yacoub, E., for the WU-Minn HCP Consortium, 2013. Pushing spatial and temporal resolution for functional and diffusion MRI in the Human Connectome Project. NeuroImage 80, 80–104.

Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E.J., Yacoub, E., Ugurbil, K., for the WU-Minn HCP Consortium, 2013. The WU-Minn Human Connectome Project: An overview. NeuroImage 80, 62–79.

Vercauteren, T., Pennec, X., Perchant, A., Ayache, N., 2009. Diffeomorphic demons: Efficient non-parametric image registration. NeuroImage 45, S61–S72.

Vos, S.B., Viergever, M.A., Leemans, A., 2012. Tract coherence imaging (TCI): Quantifying the intra-voxel fiber tract heterogeneity, in: Proc. Intl. Soc. Mag. Reson. Med., p. 3586.

Westin, C.F., Peled, S., Gudbjartsson, H., Kikinis, R., Jolesz, F.A., 1997. Geometrical diffusion measures for MRI from tensor basis analysis, in: Proc. Intl. Soc. Mag. Reson. Med., p. 1742.

Willats, L., Raffelt, D., Smith, R.E., Tournier, J.D., Connelly, A., Calamante, F., 2012. Within subject reproducibility and between subject variability of super-resolution track-weighted imaging, in: Proc. Intl. Soc. Mag. Reson. Med., p. 1918.

Yap, P.T., Chen, Y., An, H., Gilmore, J.H., Lin, W., Shen, D., 2010. Non-parametric deformable registration of high angular resolution diffusion data using diffusion profile statistics, in: Proc. Intl. Soc. Mag. Reson. Med., p. 3968.

Yap, P.T., Shen, D., 2013. Spatial transformation of DWI data using non-negative sparse representation. IEEE Trans. Med. Imag. 31, 2035–2049.

Zhang, H., Schneider, T., Wheeler-Kingshott, C.A., Alexander, D.C., 2012. NODDI: Practical in vivo neurite orientation dispersion and density imaging of the human brain. NeuroImage 61, 1000–1016.